asis&t

2001
SIG USE
Research Symposium on

Effective Methods for Studying Information Seeking and Use

asis&t

2001 SIG USE
Research Symposium on
Effective Methods for Studying
Information Seeking and Use

sponsored by the
Special Interest Group on Information Needs, Seeking, and Use (SIG USE)
of the
American Society for Information Science and Technology (ASIST)

Organizers:

Barbara M. Wildemuth
School of Information and Library Science
University of North Carolina at Chapel Hill

Ann Peterson Bishop
Graduate School of Library and Information Science
University of Illinois, Urbana-Champaign

Ruth Palmquist
Graduate School of Library and Information Science
Dominican University

Held at the J. W. Marriott Hotel, Washington, DC,
November 3, 2001

Acknowledgements

The organizing committee would like to thank the following people for contributing to the success of this symposium by participating in the reviews of the submitted papers:

Symposium Agenda

J.W. Marriott Hotel, Washington, DC
November 3, 2001

12:00   Welcome and Introductions

12:30   Contributed papers I, moderated by Ann Peterson Bishop
         Cooper
         Toms & Duff
         Hargittai

2:00   Break

2:15   Contributed papers II, moderated by Ruth Palmquist
         Heidorn, Mehra & Lokhaiser
         Gorman, Lavelle, Delcambre & Maier
         Mehra, Bishop, Bazzell & Smith

3:45   Break

4:15   Organize for roundtable discussions
         Card sorting and other categorization methods
         Direct observation of online searching
         Direct observation of group task performance
         Diaries
         Participatory action research

4:30   Roundtable discussions

6:00   Wrap-up, sharing comments from the roundtables

Table of Contents

Preface

The ASIST Special Interest Group on Information Needs, Seeking, and Use (SIG USE) was chartered in May 1999. SIG/USE members are concerned with the activities, both behavioral and cognitive, of people who are interacting with information. These activities include: recognizing information needs, seeking information that will address those needs, exploring information sources present in one's context/situation, retrieving information from available information sources, communicating and collaborating with others concerning an information need or information resources, using information, and other interactions between people and information. The SIG wishes to promote studies of human information-related behavior and provision of information services, and to encourage the application of the study results to the design of information systems and services.

The 2001 SIG USE Research Symposium was planned to address the SIG's goal of promoting studies of human information-related behavior by focusing on the research methods that can most effectively be used to study information seeking, information use, and other human information behaviors. The symposium will include the presentation of six refereed papers, plus small-group discussion of the strengths and weaknesses of several research methods currently in use.

It is hoped that this will be the first of many symposia sponsored by SIG USE, resulting in the overall strengthening of the research efforts addressed to understanding people's information needs, seeking, and use.

Barbara M. Wildemuth
Chair, Symposium Organizing Committee

Methodology for a project examining cognitive categories for library information in young children

Linda C. Cooper
Pratt Institute, Brooklyn, New York

Abstract

This paper presents an overview of some of the methodology used in a project that sought to examine children's understanding of library information and how those perspectives change in the first five years of formal schooling. Since our understanding of information is reflected in the manner in which we categorize that information (Schutz & Luckmann, 1973), the manner in which children categorize library information should provide insight regarding their perspective of the library. Data collection using group dialog, visual imagery, narrative, cooperative learning techniques, and hands on manipulatives is described for one session of a project in which children used induction to form concepts related to knowledge organization in a hypothetical library. Analysis for this session included use of hierarchical clustering and multi-dimensional scaling to examine and compare children's constructions for qualitative differences on several grade levels. Following the description of data collection methods and analysis, a discussion focuses on the reasons for using these particular methods of data collection with a child population.

INTRODUCTION

This paper presents an overview of some the methodology used in a dissertation project conducted while at Rutgers University. The larger project was concerned with the change from a personal to a cultural typification of information that a user must experience in order to successfully interact with the library. To examine this change, a study was conducted which examined the nature of and differences in the typifications of library information of children at five different grade levels in elementary school. Participants in this project were children between the ages of 5 and 9 years old. A detailed report on the project including address of theoretical considerations, results, discussion, and implications is being prepared for publication, however, since the methodology may be of particular interest as a novel approach to data collection and analysis, portions of it are presented here. Therefore, while the project had five sessions, this paper will focus on data collection and analysis of Session 3. Session 2 will be briefly described since its conduct impacts on Session 3. Mention will also be made of Session 1 since its results are used by the children in the sessions that follow it.

The goal of this project was to view the library collection from a child's perspective and to learn whether that perspective changes from personal to cultural in nature during the first years of formal schooling and regular visits to the school library media center. Since our understanding of information is reflected in the manner in which we classify, or *typify,* that information (Schutz &

Luckmann, 1973), in order to view the library collection from a child's perspective, children were invited, as described below, to shelve (i.e. classify) terms representative of library books and then to name those categories. The resulting shelf categories help us to see library information from a child's perspective. Research done by Hilda Taba describes induction as a method of concept formation (Joyce & Weil, 1996). Ideas are grouped into categories and those categories are given names, thus forming concepts. This is the method of understanding children's thinking about the library that is used in the study presented here. By constructing categories of information in this manner, children were building their idea of a library from the bottom up. They began with ideas regarding the books that a library should contain. Then they grouped those books into shelf categories and named those categories. Those shelf categories are the larger information concepts that the children envision as contained in their library.

RELATED LITERATURE

Because the purpose of this paper is to report on the methods used with a particular population, that of children, the review of related literature is here confined to studies done with children and some systems developed for children that attempt to address their typifications. Literature in areas of cognitive sociology, developmental psychology, and theories of classification pertaining to the larger project will be addressed in the separate report on that project.

Information systems currently in place are usually designed for adults and are system based in that they reflect the typifications of the library and the larger culture rather than those of the user. Research in the area of developmental psychology has shown that children think differently from adults (Carey, 1985; Seigler, 1996, 1998). If the information systems currently in place have been designed for adults, it is unlikely that they will support the information behavior of children. In addition to the problems which face adults when searching for information, children must also deal with problems related to reading and language expertise, life experience, and developmental issues such as fine motor coordination, scanning, etc. (Busey & Doerr, 1993; Eaton, 1989; Edmonds et. al, 1990; Kuhlthau, 1988; Moore & St. George, 1991; Pejtersen, 1986; Solomon, 1993). Children at the age of the participants in the present study tend to view the world from their own perspective (Piaget & Inhelder, 1969), which does not bode well for the use of a culturally typified classification system.

The most notable studies in the area of children's classification behavior have been done by Borgman (Borgman et. al 1989, 1995) and Pejtersen (1986, 1992). A computerized catalog for children was developed at the University of California by Borgman et al. (1995) to support classroom science instruction. The Science Library Catalog, developed in conjunction with Project SEED, took shape based on earlier studies (Borgman et al., 1989) conducted with children in grades three through six in which they were asked to sort words into categories to help design a computer system for libraries. The words given to the children to sort, plant terms in one experiment and animal terms in another experiment, were chosen from glossaries of the grade level science textbooks used in the children's school. Children did not always have adequate knowledge of the terminology used in this project even though the terms were chosen from grade level glossaries. Results of this study indicate that "children can think in categorical

terms about the domain of the database provided they are given vocabulary terms that they understand" (Borgman et al., 1989; p. 86).

Pejtersen (1986, 1992) designed and tested a database for fiction based on an analysis of children's information seeking behavior. Using information gathered in user-librarian interactions in a library setting, Pejtersen sought to "develop a classification model that reflect[s] the children's value perception when selecting documents and which could serve both as a model for the concept analysis of documents and for the design of a computer based retrieval system with combinatorial retrieval access to several dimensions of document features. To achieve this, the children's value criteria should be known through their formulation of needs and their characterization of document contents in real life searches...The main difference between the adults' search behavior and the children's search behavior is found in the Subject-matter dimension and in the Accessibility dimension. Not surprisingly, the Accessibility of documents is by far the most frequently used value criterion" (Pejtersen, 1986; p. 128).

In the project described below, the sort terms are not supplied for the children. In this project, the children generate suggestions for the items that they will sort. This gives us an idea of what type of information they typify, or envision, as belonging in a library. While choosing their own sort terms does not guarantee that the children understand the terms in the same way that adults do (Vygotsky, 1986, p.133; Wertsch & Stone, 1985, p.169), choosing their own terms does imply that the children have some kind of understanding or knowledge of those terms. The manner in which they categorize these terms and names they give to the categories they construct should lend some insight into their understanding of how information in the library is interrelated and the criteria upon which they base these relationships. This, in turn, may suggest system design more supportive of access both for children and perhaps other users who are not familiar with the cultural typifications of library organization. Freedom of intellectual pursuit without the filter of another person's or culture's typifications supports the realization of individual potential.

Several lesser know efforts at accommodating children's typifications of library information bear mentioning. The Franklin Lakes Public Library (Franklin Lakes, New Jersey) Preschooler Door to Learning Project developed The Catalog for Preschoolers (Crocker, 1994) so that preschool children might search for and find books without adult assistance. Most often, very young children must depend on an adult intermediary to interpret their information request and access the material for them. This system used icons so that children who were not yet able to read could search unassisted. Developers worked on the premise that even though they may not be able to articulate their information need, young children understand enough about their need to recognize and use visual cues to find information.

The Waltham Forest School in the London Borough of Waltham Forest developed a system that they feel is more in keeping with the vocabulary used in the average child's school curriculum (Tyerman, 1989). The library collection is color coded and broken down into 20 categories based on where the children's librarian felt children would look for information. Vocabulary used is based on that which is used in the school curriculum.

The Common School, a private K-6 school in Amherst, Massachusetts uses a color-coded thematically arranged system which they feel is in keeping with the developmental level of their

user population (Cooper, 1997). Their collection is broken down into Fiction; Biography; Countries and Cultures; Myth, Folk and Fairy; Beginning Readers; Language; Arts; Science; and Picture Books rather than the Dewey system used in most children's libraries. Books are kept together by subject, in keeping with literature in developmental psychology that indicates that young children tend to categorize thematically (Carey, 1985; Keil, 1989; Markman & Hutchinson, 1984).

These are just three local efforts and doubtless there are more. These libraries have made a real effort to make information accessible to their own population of children based on their typifications. However, in two out of three instances, the children's typifications have been interpreted by adults. The collections have been arranged based on what adults think children think. Session 3, described below, is designed to examine where children would place information if they were organizing the library themselves according to their own typifications.

METHODOLOGY

Background

This study was conducted in a primary level public school (grades k-2) and an intermediate level public school (grades 3-4) in a middle class suburban area of New York. Children in both schools visit the school library media center once a week for a lesson in information seeking or literature appreciation as well as the opportunity to borrow a book to take home for the week. The primary school in which the study took place serves 413 children - 119 kindergarten children, 154 first grade children, and 140 second grade children. There are no special education classes in this building. The school does offer supplementary classes for children for whom English is a second language and a resource room to assist children needing extra help in reading and math. Ethnic diversity is somewhat limited and the majority of children are Caucasian. Economic background of children is somewhat, though not largely, diverse. Male/female ratio of students is approximately equal. All children in this school participated in this project and each session lasted approximately 25 minutes.

The intermediate school in which the study was conducted is in the same school district and serves children in the third through fifth grades. Ethnic diversity, male/female ratio of students, and socio-economic background of students is similar to that described for the primary school. This school does serve special education students and two of the fourth grade classes participating in this study are conducted in an inclusion fashion during scheduled library visits. Approximately fifty third graders and fifty fourth graders participated in this study. Approximately ten of these fourth graders were from a special education class. Children were permitted to work in groups of their choice so that a work group may or may not be composed of combined classes. Each session of the project in this school lasted approximately 20 minutes. Students in first, second, third, and fourth grades did the same type of exercise, however, the terms they sorted were different for each grade level. Kindergarten exercises differed, as described below, since many of the kindergarten children cannot yet read. The researcher is the school library media specialist (slms) in both situations.

Definition of terms

category - all the terms that a group of children feel belong together as books on a library shelf

class - those children who share a common classroom and teacher for a school year and visit the library together once a week

group - approximately four children who are working together throughout this project; each class of children had approximately six groups

term - a word (i.e. cat) or words (i.e. Spice Girls) suggested by children in a class during Session 1 as being something important to include information about in a library and sorted in Session 3 as though it were a book on a library shelf

label - the name that a group of children gave a category of terms, a shelf label

library information - information found within a library as distinguished from everyday information

Session 1

In Session 1, children were asked to visualize an empty library for which they would choose the books. Each class was asked to offer suggestions for what they felt were the most important things include in a library. These suggestions were written on a large board so the children could see them. The purpose of this session was to gain insight into what the children typified as important in their concept of a library. A detailed description of Session 1, including methodology, results, and discussion will be the topic of a separate paper.

The suggestions from each grade level in this session supplied a pool of terms from which the slms/researcher selected the sort terms to be used in Session 3. These terms were selected based on two criteria. In general, terms selected for inclusion were those suggested most often by the grade level. In addition, suggested terms were also included if it appeared that they would require deeper consideration in order to categorize them. For example, *Harry the Dirty Dog,* was not one of the *most* often suggested terms by Grade 1, but it was included as a sort term nevertheless in order to discern where children would categorize it - with books about dogs, in a broad category with other animals, with picture books, with pets or someplace else.

The following terms suggested by Grade 1 in this session were selected by the slms/researcher and later presented to them to sort in Session 3 described below: Amelia Bedelia, American Girl, Animorphs, Arthur, atlas, Baby-Sitters Club, Back Street Boys, baseball, bats, birds, cats, cheetahs, Clifford, Curious George, dictionary, dinosaurs, dogs, dolphins, Dr. Seuss, dragons, elephants, fish, football, frogs, Furby, giraffes, Godzilla, Harry the Dirty Dog, Henry & Mudge, hockey, I Spy, knights, komodo dragons, Lassie Come Home, Lyle Crocodile, magic, monkeys, mouse, penguins, piranhas, planets, Pokemon, polar bears, presidents, princesses, Rainbow Fish, Robert Munsch, Rugrats, seals, sharks, snakes, Spice Girls, Star Wars, tarantulas, tigers, Tommie De Paola, whales.

Session 2

During Session 2, we imagined that the terms selected in Session 1 were books that had been delivered and were now unloaded from the delivery truck and in a pile on the library floor. The slms/researcher asked the children if it would be easy to find a desired book in this pile. When the children responded in the negative, they were then asked what needed to be done to make it easier to find a book they might want. A class discussion led to the suggestion that the books needed to be sorted on to shelves. We talked about sorting, a concept that is addressed in Kindergarten and one with which the children were readily familiar. Children offered various examples of how things might be sorted, for example by size, color, or shape. At this time, we did not yet talk about sorting books but rather other things such as blocks or beads. These are the types of things that they are most accustomed to sorting. We read either <u>The Button Box</u> (Reid, 1990) or <u>A String of Beads</u> (Reid, 1997). While it was important to review the concept of sorting at this time, it was also important not to suggest ways that books might be sorted. The purpose was not to impose typifications upon the children, but rather to allow them to respond with their own typifications.

Then we began to address the problem of sorting the children's hypothetical book delivery. Children were invited to sort some of their responses from Session 1 so that the books would be in an order conducive to easy retrieval. This was a whole class session to prepare children for the following session in which Grades 1 - 4 broke up into smaller groups and worked on their own. Terms used to model sorting in this session came from the pool of terms that the children mentioned in Session 1 as important things to have in a library, but they were not reused in Session 3.   For Grades 1 - 4, the slms/researcher wrote approximately 20 terms that each grade level had suggested on separate index cards. She held up the first card for the class to see and attached it to the white board. Then she held up the second card and asked whether this book should be on the same shelf as the first or on a different shelf. If the shelving was the same, the card was tacked to the board directly under the first card. If the shelving was different, the card was tacked to a removed area of the board. The third book could be placed on either of the first two shelves or on a third shelf. In this way, we proceeded though these practice terms and the end result was several groups containing terms representative of books that the children thought belonged together on a library shelf. The slms/researcher then asked the children to name each shelf. For example, a shelf on which they put Cats, Rabbits, and Lassie might possibly be named Animals. This method is similar to the card sort method used in web site development (Nielsen, 1993, Bernard, 2000).

Since children in Kindergarten were not able to read well enough to sort words on index cards, they did not have a session like the one described above. Their sorting exercise is described below in Session 3.

Session 3

Children in grades 1 - 4 broke up into groups of their own choosing of approximately four. Each group received an envelope containing the terms indicated above suggested by their grade in Session 1. Each term was written on a small slip of paper approximately 1/2"x2". Each group

also received one 12"x18" piece of paper and several glue sticks. The groups were instructed to sort the terms into categories of books that would "go together" on a library shelf in the same manner that we had practiced in Session 2. The large piece of paper would serve as their work board. When they were satisfied with their shelf categories, they were to glue the terms in place and then label each category with a name that described the group of books they placed on that shelf. Not all groups finished sorting the entire envelope of terms. In general, upper grades finished sorting more terms.

Each Kindergarten class was addressed as a single group since many of those children were not yet able to read and may have lacked the fine motor coordination needed to work with the slips of paper as well as the social skills needed for group decision making. Books whose covers depicted the Kindergarten's most frequently mentioned terms from in Session 1 were displayed to the class. These were chosen by the slms/researcher from the shelves of the school library. For example, if the term named by the children was Cat, then a book with a photograph of a cat on the cover was chosen. If the children named Clifford as important, then a book with a picture of Clifford on the cover was chosen. The slms/researcher assisted the class in physically sorting the books into groups that they felt belonged together. Then they named the groups and the slms/researcher noted down the contents and names.

In these sessions children critically considered what should be in a library and how the material in a library should be categorized. Children were instructed to come to a group decision. This was not always easy for them to do. Indeed, there were heated arguments regarding the categorization of information. The children were instructed by the slms/researcher that the opinion of the majority of the group should direct their final decision. The group work, therefore, provided an excellent opportunity for cooperative learning among the participants. Data were collected in the form of the field notes and children's handwork (words pasted on paper). Table 1 presents an example of the efforts of one group of four children from Grade 1 in Session 3.

Table 1. Category contents and labels constructed
by one group of Grade 1 children during Session 3

Category label	Category contents (in order of placement)
Baby siters (sitters)	Baby Sitter's Club
Aouthers (authors)	Robert Munch
Reptiels (reptiles)	dragons, Godzilla, komodo dragons, snakes
Lizirds (lizards)	dinosaurs, Lyle Crocodile
Spoa (sports)	baseball, hockey
Cartoun (cartoon)	Rugrats
Mesturey (mystery)	I Spy
Adamles (animals)	penguins, dogs, Harry the Dirty Dog, Furby, Lassie Come Home, giraffes, birds, polar bears, tigers, monkeys, sharks, piranhas, seals
Ficktchain (fiction)	Pokemon, Arthur
Pepole (people)	presidents, knights, Spice Girls, Amelia Bedelia, princesses

Analysis of Session 3

Approximately 500 children participated in this project. In Session 3, children constructed 995 categories altogether and articulated 353 category names (there are more categories than names because some names were articulated more than once). In order to discern which terms children associated with each other, a method needed to be devised to find co-occurrence of terms within the categories constructed on each grade level.

The member terms of each category and the label assigned to that category by the group of children who constructed that category were entered into spreadsheets. This was done separately for each grade level so that later a comparison might be made. The most frequently named categories and their contents on each grade level were calculated. With the help of Dr. Ronald Rice of Rutgers University, the spreadsheets were then entered into UCINET (Analytic Technologies, 1996). This program was used to generate a co-occurrence matrix for each grade level that indicated the number of times any of the terms occurred together in a category with any other of those terms. Based on these co-occurrence matrices, UCINET also generated multidimensional scaling and hierarchical clustering output for each grade level.

Cluster analysis is a statistical procedure that organizes a data set of entities into "relatively homogenous groups" (Aldenderfer & Blashfield, 1984, p. 7). Those terms sorted by the children that co-occurred in categories the most are represented in these outputs as having the highest degree of clustering. Terms that co-occur less often are represented as having proportionally a smaller degree of clustering. These multidimensional scalings and hierarchical clusterings, along with the data regarding the most frequently named categories and their contents provided a starting point for the analysis of the data. A brief description of the hierarchical clustering and multidimensional scaling (MDS) analysis is presented here.

UCINET generated hierarchical clustering output for each grade level. The hierarchical clustering output has a graph-like appearance. In these graphic representations of the data, major clusters were clearly separated thus making interpretation of the degree of co-occurrence of terms fairly straightforward. The output from the MDS is a graphic representation similar to a map showing proximities between objects (Kruskal & Wish, 1978). In this case, the objects are the terms that the children sorted. MDS provided an even more visual representation of the proximities or clusters of terms presented in a hierarchical clustering. The terms that tend to co-occur more often are represented in the MDS as being graphically closer to each other. The MDS output has a more map-like appearance. In this study, the children were given terms to sort into categories representative of books that they felt belonged together on a library shelf. The resulting MDS represents, therefore, a sort of cognitive map of how these children perceive the information in a library. Examination of the MDSs in this study involves not only note of which terms are clustered together but also how these clusters relate to each other in dimensions from left to right and top to bottom of the MDS as a whole.

Both two and three-dimensional MDS representations of the co-occurrence matrices based on the data were generated for each grade level in Session 3. Both were examined to glean information regarding children's typifications of the library on each grade level. After verifying that the larger patterns represented by both dimensions remained the same in each instance, the two

dimensional MDS coupled with hierarchical clustering were used as the starting point for discussion since two dimensional output is more easily addressed than three dimensional output (Trochim, 1989; Kruskal & Wish, 1978; Borgatti, 1997). The output from Grade 1 will serve as an example to explain how the analysis proceeded.
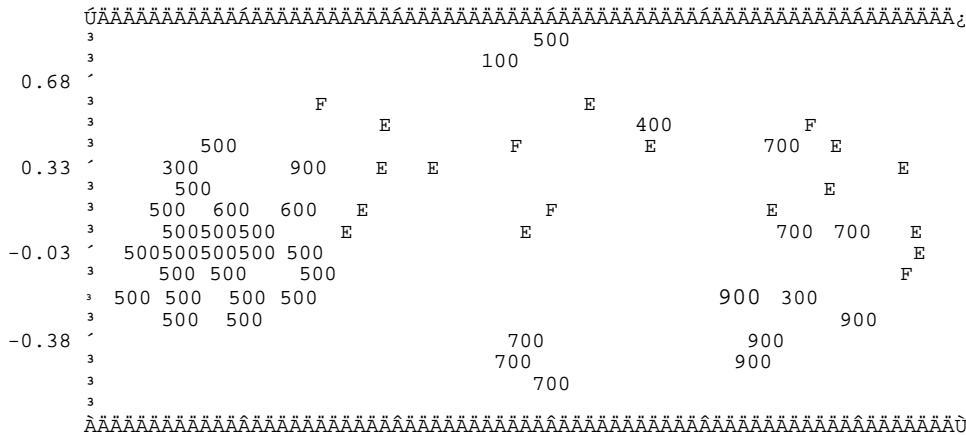
Words in the MDS outputs are not always spelled out completely, thus making the concept map more difficult to read. Also, not all terms sorted are visible in the two-dimensional output. Therefore, each term was fully spelled and plotted on the two-dimensional MDS grid according to its coordinate locations. This gave a more understandable picture of how the terms are conceptually located in relation to each other (see Figure 1).

Figure 1. Session 3 Grade 1 MDS

```
        ÚÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ¿
        ³                              PLANETS
        ³                         ALIENS        HENRY&MUDGE
  0.68  ´
        ³              GODZILLA
        ³                 CURIOUSGEORGE              DICTIONARY      BABYSC
        ³         DINOSAUR                LASSIECH   DRSEUSS  ISPY ARTHUR
  0.33  ´      DRAGON        ATLAS  HARRYDD CLIFFORD                    RMUNSCH
        ³      KOMODOD                                        POKEMON
        ³    BATS   CATS  DOGS RAINBOWFISH      ANIMORPHS       RUGRATS
        ³  GIRAFFE  CHEETAH POLARBEAR LYLEC    FURBY          MAGIC STARWARS AMELI
 -0.03  ´ ELEPHANT TIGER MONKEY BIRDS FISH                             TDEPAOL
        ³   MOUSE FROG     DOLPHIN                                  Amrgirl
        ³ TARANTULA PENGUINSHARK SEAL WHALE          PRESIDENTS PRINCESSESS
        ³  PIRANHA  SNAKES                                    KNIGHTS
 -0.38  ´                         FOOTBALL         BACKSB
        ³                         BASEBALL         SPICEGIRLS
        ³                          HOCKEY
        ³
        ÀÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ
```

In order to compare this with the manner in which the library classifies the same terms, a broad Dewey number was assigned to each term sorted according to the Dewey Decimal method of Classification (Dewey, 1990). For example, the term Dinosaurs was assigned the number 500; a book about Clifford was assigned an F for fiction. In this way, the children's shelving of their imaginary books could be compared to the manner in which the library would shelve those books. Recall the instructions that the children received were to consider each term a book to be placed on shelves with others that belong together to facilitate easy access. These Dewey classifications were then plotted on the MDS grid in a manner corresponding to the children's term placement. This gave a graphic representation of the children's term placement compared to a Dewey number. It compared the children's typification of library information to the cultural typifications used by the library to organize its information. It showed, for example, if children shelved Curious George separately or in close proximity to books about gorillas (see Figure 2).

Figure 2. Session 3 Grade 1 Dewey numbers plotted

```
ÚÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ¿
      ³                                      500
      ³                                    100
 0.68 ´
      ³                    F                  E
      ³                        E                  400           F
      ³          500                   F              E       700   E
 0.33 ´      300           900     E    E                                      E
      ³       500                                                         E
      ³     500  600   600    E              F             E
      ³      500500500     E              E                    700  700     E
-0.03 ´   500500500500 500                                                      E
      ³       500 500      500                                                F
      ³   500 500   500 500                          900  300
      ³       500   500                                          900
-0.38 ´                           700              900
      ³                           700              900
      ³                             700
      ³
ÀÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÙ
```

Examination of these MDSs shows that for children in Grade 1 an important aspect of classification is the definition between person and animal. Fiction and non-fiction about animals may be clustered. For example, Curious George, Harry the Dirty Dog, Rainbow Fish, and Clifford are placed on the side of the MDS that holds books about real animals. However, books about animals and people, whether fiction or non-fiction, are clustered quite separately. The left right dimensionality moves from animal terms to people terms. The up down dimensionality moves from fantasy based terms (aliens, Godzilla, Dr. Seuss) to reality based terms (baseball, dolphins, Spice Girls). It is also interesting to note that, according to this interpretation, dictionaries hold less reality to Grade 1 children than Furby. Upon consideration, this is not surprising since Furby, in addition to being a cartoon character, may also be a toy, a concrete object they may own. Even as a cartoon, it occupies a larger place in the life-world of a child than does a dictionary. Animorphs (a series about children who have the ability to morph into animals) is placed in the mid-point between animals and people, real and not real - not far from Furby.

The hierarchical clustering output indicates major groups of co-occurring terms that can be broken down into smaller and smaller clusters of co-occurrence (see Figure 4). These clusters of terms were compared to the contents of the most frequently named categories calculated earlier from spreadsheets in order to see what category names the children had given them. This gave a sense of what the children felt was the most salient quality of the term groups, the criteria on which they were basing their categories. Using this information, the children's category names were placed on the MDS grid so that these names corresponded spatially to the positioning of their term groups. This gave a conceptual map of the children's category names (see Figure 3).

Figure 3. Session 3 Grade 1 category labels plotted

```
     ÚÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ¿
     ³                        PLANETS
     ³                          SPACE
0.68 ´
     ³                      DOGS
     ³        LIZARDS    ANIMALS   BOOKS              BOOKS
     ³           DINOSAURS                         STORY
0.33 ´                                        CARTOONS   BOOKS
     ³        CATS                             TOYS      AUTHORS
     ³          DOGS                  ANIMALS
     ³        ANIMALS              TOYS
-0.03´
     ³                                              PEOPLE
     ³
     ³
-0.38´                                        SING
     ³                   SPORTS            MUSIC
     ³
     ³
     ÀÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÙ
```

Figure 4. Session 3 Grade 1 hierarchical clustering.

```
JOHNSON'S HIERARCHICAL CLUSTERING
ÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ

Input dataset:          C:\COOPER\G1SQ
Method:                 AVERAGE
Type of Data:           Similarities

HIERARCHICAL CLUSTERING
```

(dendrogram output)

Examination of the hierarchical clustering chart reflects points similar to those observed on the MDS. There are several main divisions in the co-occurrence of terms. The largest and most pronounced contains animals and Animals is the label which children most often gave this category. The other side of the hierarchical clustering chart is divided into parts reflective of a more social or people oriented domain. The largest of these contains terms which could be considered easy or picture book fiction and the children most often labeled this category Books or Cartoons or Story. Most of the terms on the far left side and the far right side of the chart are real things. The terms clustering towards the middle of the chart are things that are less real. Again, we see that qualities of animal v people and real v not real are considerations with Grade 1 in the organization of their hypothetical library.

All of this is in keeping with the information regarding term clustering gleaned from examination of the MDS so that the outputs support similar conclusions. According to both the hierarchical clustering and the MDS, while Grade 1 considers the quality of Animal a salient one for both Clifford and Dog, there is beginning to be more discretion regarding their placement together on the same library shelf than the Kindergarten exhibited. While they occur on the animal side of the MDS, Clifford is placed on the edge of the animal cluster rather more towards the fantasy dimension. We can see on the hierarchical clustering chart that Clifford is placed on the midpoint between Animals and People. There is more awareness of reality v fantasy than there was in Kindergarten. While Clifford is usually put on a shelf called Animals or Dogs, he now also appears fairly often on shelves called Story, Cartoons, and Books.

DISCUSSION

Scope and limitations of study

This study was conducted in the context of a real school setting, as part of the children's regular school routine as opposed to a laboratory or contrived situation. An effort was made to conduct sessions in as normal a manner as possible so that the children's responses would not be contrived or stilted. All the variables that normally impact on a regular school library situation were, therefore, in effect here. While Krathwohl warns that such qualitative methods may result in "multiple" or "context-dependent" explanations (Krathwohl, 1993, p. 311), certainly *everything* is context dependent and, in fact, anything taken out of context is without meaning. A study conducted in a non-normal setting would result in data affected by that non-normal context.

Children in this study were encouraged to construct arrangements of information that made good sense to them, not necessarily arrangements that they have observed to be in place in the library. Adult assistance was not encouraged. However, because the average age of the child participant was seven years, it was felt that having the children work individually would be too difficult for them. To ask twenty-four children to work independently sorting, gluing, and naming 50 terms in 25 minutes would have been a logistical nightmare for the slms/researcher and an overwhelming task for the children. Working in groups of four gave them the support that they needed to complete the project. The resulting categories were built by a small group of children rather than one individual child but the categories were constructed by children without adult intervention. It is felt, therefore, that the positive aspects of small group work, in this instance, are more

important than considerations regarding group dynamics and that the resulting data are representative of the information choices of the 'average' participant.

Since this study was conducted in a qualitative manner with a particular population, the findings reflect only the behavior of children in this study. However, since the participants numbered more than an entire school, it is felt that findings regarding children's typification of library information may have some broader interpretation. The reader is reminded, however, that the participating children are of a broad common background and their understandings certainly will differ from children raised in a different culture. This caution, however, applies to a myriad of other studies and is not unique to the study described here. This study did not address impact of group dynamics, gender of participants, or presence of the school library media specialist on the children's typifications. All these variables are acknowledged and removal or control of any of them would have changed the participant's situation from 'normal' to 'not normal.' That type of situation would have been contrary to the research perspective stated above.

Concluding remarks

The visual presentation of output for the hierarchical clusterings and MDSs contributed greatly to the manner in which the data were interpreted. Examination of the hierarchical clustering and MDS output for each grade level enabled the researcher to:
1. gain a sense of how each grade level cognitively classifies the information in the library;
2. compare the children's classification to those which the library uses to categorize its books;
3. learn whether and how children's classification of library information changes as they get older;
4. learn whether children's classification of library information moves towards that of the library.

Levi-Strauss describes the 'bricoleur' examining a set of materials who must "consider or reconsider what it contains and, finally and above all, ... engage in a sort of dialogue with it and, before choosing between them, to index the possible answers that the whole set can offer to his problem" (Levi-Strauss, 1966; p.18). Children constructing categories in the present study may be considered bricoleurs. They built shelf categories using terms "the possible combinations of which are restricted by the fact that they are drawn from the language where they already possess a sense which sets a limit on their freedom of maneuver. And the decision as to what to put in each place also depends on the possibility of putting a different element there instead, so that each choice which is made will involve a complete reorganization of the structure, which will never be the same as one vaguely imagined nor as some other which might have peen preferred to it" (Levi-Strauss, 1966; p.19).

Much of the analysis involved sorting data which the children had generated by sorting, as well. "[T]he signified changes into the signifying..." (Levi-Strauss, 1966; p. 21). That is, children were given the information that they generated during Session 1 to categorize in order that the researcher might gain an insight into their understanding of it. Their responses to this exercise, in

turn, were categorized in the analysis. This analysis, in a sense, is again in keeping with Levi-Strauss' description of the 'bricoleur.'

The methodology described here evolved as the project progressed more than it was pre-designed. A pilot project performed the previous year helped to shape the data collection methods used here. The method of analysis sprang from the data that the children generated. An effort was made to use collection methods that were in keeping with the participating population and which would give something to these children as well as elicit the information from them that was needed for the study. In these sessions children critically considered what should be in a library, how that information might be made accessible, and how the material in a library should be categorized. Discussion within groups involved decisions about information organization and highlighted the different manners in which information can be organized in a library. Thus, the children moved to the upper level of Bloom's Taxonomy (Bloom et al., 1956) through analysis of the terms which they had generated in Session 1 for meaning and their relation to other terms, and through synthesis of these results to build their own organization of information. Based on the slms/researcher's observations of the children during the project, as well as their comments during a debriefing which took place after Session 5, it is observed that the children came away from the project with a deeper understanding of knowledge organization and the problems involved with organizing information for a community of users as well as a valuable exercise in cooperative learning. Visual imagery, narrative, and working with their hands contributed to the children's enjoyment, understanding, and learning experience.

One of the most important considerations in data collection was to avoid directing the children's responses. Children were not asked to learn someone else's way of thinking; rather, the purpose of the study was to gain insight into their way of thinking. Children are accustomed to getting information from adults, particularly teachers. In this project, the researcher was the children's school library media specialist, making her a participant observer. To this end, great care was taken to avoid directing the children's responses. All of the terms they suggested, the categories they constructed, and the labels they articulated were accepted. In this way, and through the children's own constructions, the concept of the library and its organization was examined through the eyes of the child user.

ACKNOWLEDGEMENTS

REFERENCES

Aldenderfer, M. & Blashfield, R. (1984). *Cluster Analysis*. California: Sage Publications.
Bloom, B., Englehart, M.B., Furst, E.J., Hill, W.H., & Krathwohl, D.R. (1956). *A taxonomy of educational objectives: Handbook I. The cognitive domain.* New York: David McKay.

Bernard, M. (2000). *Constructing user-centered websites: The early design phases of small to medium sites*. [Online] Available: http://156.26.13/newsurl/usabilitynews/2W/webdesign.htm

Borgatti, S. (1997). *Multidimensional scaling*. [Online] Available: http://www.analytictech.com/networks/hiclus.htm

Borgman, C., Chignell, M., & Valdez, F. (1989). Designing an information retrieval interface based on children's categorization of knowledge: A pilot study. In J. Katzer, G. Newby (Eds.), *Managing Information and Technology: Proceedings of the 52nd Annual Meeting of the American Society for Information Science 2.* October 29-November 2, 1989, Washington, DC. Medford, NJ: Learned Information, 81-94.

Borgman, C., Hirsh, S., & Walter, V. (1995). Children's searching behavior on browsing and keyword online catalogs: The science library catalog project. *Journal of the American Society for Information Science*, 46, 663-694.

Busey, P. & Doerr, T. (1993). Kid's catalog: An information retrieval system for children. *Journal of Youth Services in Libraries*, 7, 77-84.

Carey, S. (1985). *Conceptual change in childhood*. Massachusetts: MIT Press.

Cooper, L. (1997). The retrieval of information in an elementary school library media center: An alternative method of classification in The Common School Library, Amherst, Massachusetts. *Public & Access Services Quarterly*, 2, 1-24.

Crocker, S. (1994). The catalog for preschoolers. Presentation as part of the symposium *Catalogs For Kids: Organization, Access, and Interfaces at the Congress For Librarians*, St. John's University, Division of Library and Information Science, Jamaica, N.Y., 21 February 1994.

Dewey, M. (1990). *Abridged Dewey Decimal Classification and relative index. Ed. 12*. John P. Comaromi, Ed. Albany: Forest Press.

Eaton, G. (1989). What the public children's librarian needs to know about locational skills instruction in elementary schools. *Journal of Youth Services in Libraries*, 2, 357-366.

Edmonds, L., Moore, P., & Balcom, K. (1990). The effectiveness of an online catalog. *School Library Journal*, 36, 28-32.

Keil, F. (1989). *Concepts, kinds, and cognitive development.* Massachusetts: MIT Press.

Krathwohl, D. (1993). *Methods of educational and social science research: An integrated approach.* New York: Longman.

Kruskal, J. & Wish, M. (1978), *Multidimensional Scaling.* California: Sage Publications.

Kuhlthau, C. (1988). Meeting the information needs of children and young adults: Basing library media programs on developmental states. *Journal of Youth Services in Libraries*, Fall, 51-57.

Joyce, B. & Weil, M. (1996). *Models of teaching.* Needham Heights, Mass.: Simon & Schuster.

Levi-Strauss, C. (1966). *The savage mind.* Chicago: The University of Chicago Press.

Markman, E. & Hutchinson, J. (1984). Children's sensitivity to constraints on word meaning: Taxonomic versus thematic relations. *Cognitive Psychology, 16,* p. 1-27.

Moore P. & St. George, A. (1991). Children as information seekers: The cognitive demands of books and library systems. *School Library Media Quarterly,* 19, 161-68.

Nielsen, J. (1993). *Usability engineering.* San Francisco, California: Morgan Kaufmann Publishers, Inc.

Pejtersen, A.M. (1986). Design and test of a database for fiction based on an analysis of children's search behaviour. In P. Ingwersen, L. Kajberg, & A.M. Pejtersen (Eds.), *Information Technology and Information Use: Towards a Unified View of Information and Information Technology* (pp. 125-146). London: Taylor Graham.

Pejtersen, A.M. (1992). New model for multimedia interfaces to online public access catalogues. *The Electronic Library, 10 (6),* 359-366.

Piaget, J. & Inhelder, B. (1969). *The psychology of the child.* New York: Basic Books.

Reid, M. (1990). *The button box.* New York: Dutton Children's Books.

Reid, M. (1997). A string of beads. New York: Dutton Children's Books.

Schutz, A. & Luckmann, T. (1973). Structures of the life-world. Evanston: Northwestern University Press.

Siegler, R. (1996). *Emerging minds.* New York: Oxford University Press.

Siegler, R. (1998). *Children's thinking.* New Jersey: Prentice-Hall.

Solomon, P. (1993). Children's information retrieval behavior: A case analysis of an OPAC. *Journal of the American Society For Information Science, v. 44,* p.245-64.

Trochim, W. (1989). An introduction to concept mapping for planning and evaluation. In W. Trochim (Ed.) *A Special Issue of Evaluation and Program Planning*, 12, 1-16. [on-line] Available: http://trochim.human.cornell.edu./research/epp1.htm

Tyerman, K. (1989). Alternative arrangements: Children's non-fiction categorization in Waltham Forest. *Library Association Record*, 91, 393-94.

UCINET (1996). Analytic Technologies, 11 Ohlin Lane, Harvard, MA 01451. [Online] Available: http://electic.ss.uci.edu/~lin/ucinet.html

Vygotsky, L. (1986). Alex Kozulin (Ed.) *Thought and language*. Massachusetts: MIT Press.

Wertsch, J. & Stone, C. (1985). The concept of internalization. In Wertsch, J. (Ed.) *Culture, communication and cognition: Vygotskian perspectives*. New York: Cambridge University Press.

"I spent 1 ½ hours sifting through one large box....": Diaries as Information Behaviour of the Archives User

Elaine G. Toms and Wendy Duff
Faculty of Information Studies, University of Toronto

INTRODUCTION

The diary is the "document of life" (Allport, 1943), recounting an individual's public and private thoughts in a particular context and instant (Plummer, 1983). Each episode in a diary embodies the perspective, "I am here, and it is exactly now" (Forthergill, 1974). As such diaries hold great promise for the study of information behaviour as they capture rich detail in highly specific contexts and do so using a natural process.

We are studying how historians find what they need in an archive, as the initial phase of a larger project that investigates usable finding aids. There is little known of how users can and do cope with the complexity of archival finding aids, and only a speculative understanding of what finding aids are used for or, indeed, how they are used in the scholarly process. Much of the work to date has involved understanding what historians do when they conduct research (e.g., Case, 1991; Cole, 1998) and not how they use archives and archival finding aids within that process.

In our research, we are using multiple techniques: interviews, diaries and observation to capture what historians do on entry into an archive. Each technique serves a unique purpose in our design. The interviews provide high-level descriptions of why scholars are seeking information and subjective evaluations of various access methods and tools. The diaries provide a more specific level of detail about information seeking events, documenting the various tasks undertaken with a tool or other person. Finally observation will allow us to determine explicitly how scholars use an access tool from an objective observer perspective. Whereas interviews provide a big picture view of the scholar's working environment and observation enables us to isolate precise steps, it is the diary that should provide rich detail about use in a particular context, and like think-alouds close to the occurrence. In this paper, we examine diaries as a method of data collection and outline how we are implementing in a study of historians doing research in an archive.

PREVIOUS WORK WITH DIARIES

Diaries have long been a key source of data for biographers and literary scholars. More recently the technique has been adapted for use by psychologists, sociologists, health care researchers, market researchers and information scientists. While the former use the traditional daily personal journal, the latter use a purposeful and often structured account of a limited period or specific event in a person's life – the research diary.

Diaries are all of a distinct genre – the personal life record, but not all diaries contain the same type of information, have the same form and can be used for the same purposes. Research techniques using diaries are one of four types: the intimate journal, the memoir, the log, and the diary-interview method (Allport, 1943; Plummer, 1983):

1. The intimate journal contains chronologically ordered, free-flowing entries, often of a highly personal and intimate nature. Unlike the typical daily journal, the research version may include more entries over a limited time period.

2. The memoir is written after the event(s) and often all at once; it may capture the episodic character of the genre but from a less personal perspective. Somewhat analogous to 'tomorrow' diaries, memoirs may suffer from recall problems (Corti, 1993) and the intent to write for an audience (Elliott, 1997) rather than for oneself.

3. The log, first devised by Sorokin and Berger (1939), and on which the transaction log is based, records an impersonal account of a series of structured and usually chronological activities. The technique has been widely used in many areas to acquire specific and often quantitative details when the data cannot be objectively and accurately acquired in a scientific way. The range of studies that have used this technique is varied: time schedules (Sorokin & Berger, 1939), health practices (Verbrugge, 1980), the use of engineering and technical information sources (Allen, 1996) and book use (Beheshti, 1986) to name a few.

4. The diary-interview method first described by Zimmerman and Wieder (1977) includes a two-part procedure that combines a type of diary with intensive interview. Zimmerman and Wieder designed a structured diary to record daily interactions. Participants were prompted for who, what, where, when, how, with a lengthy probing interview that is analogical to the talk-after in user testing. In the interview, each aspect of the diary was discussed for elucidation and amplification and for checking the internal consistency of the entries. Corti (1993) claims that this is the most reliable method for obtaining information.

The key advantage to diaries is the short term between event occurrence and record of that event. A diary embodies a discontinuity that reflects an "ever-changing present" (Elliott, 1997). Unlike interviews and focus groups, the diary enables the capture of a participant's moment in time rather than reflection after the fact. It is thus less subject to memory lapses, retrospective massaging, or potential for relating an event in a different context (Verbrugge, 1980). The diary also serves a duo role; it exploits the participant as both observer and informant (Zimmerman & Wieder, 1977). It may contain both accounts of an individual's performance and activities as well as those of others with whom he/she has interacted. Unlike think-aloud protocols and observation, a diary reflects a person's communication with him/herself and not a public and unnatural exposition. Dairies provide the participants with an opportunity to records both their thoughts as well as their actions as they seek information. Diaries are an alternative to observation as they eliminate the potential for modified participant behaviour and enable data collection in situations that are not observable (Zimmerman & Wieder, 1977). Diaries generally have higher levels of reporting (Verbrugge, 1980) specific details than interviews (Silberstein & Scott, 1991) and enjoy increased reliability (Hilton, 1989; Conrath, Higgins & McClean, 1983),

particularly the diary-interview method. That said, the diary is a research instrument that is written for the researcher and with the researcher's needs in mind.

Diaries require participant cooperation to succeed. Over time participants may become sensitive to the process and may modify their behaviour (Verbrugge, 1980). Often participants suffer from first-day/week effects recording more content initially that trails off toward the end of the data collection period. One of the key problems cited is sampling (Goodall, 1994) and retention. Depending on the time frame and complexity of the task, it is likely only the most persistent individuals who will finish the job. Several researchers (e.g., Mellon, 1990; Kuhlthau, 1993) have managed to include the diary writing as part of a school course requirement, increasing the likelihood of completion, but also questioning the reliability of the entries; the diary becomes an assignment with all the baggage that the process creates. The most significant advantage of diaries – the tight coupling of event and record – may be lost if participants treat the activity more like a memoir (Davis, 1971).

Diaries have been used selectively in library and information science research (Goodall, 1994). Wood (1969) suggests that diaries are inappropriate for use studies because subjects pressed for time may document their activities after the event instead of recording their actions as they happen. Furthermore, diaries take a considerable amount of time to analyze and consequently, usually involve only small numbers of participants. Thus, not everyone agrees that diaries are appropriate.

Allen (1996) used logs to better understand the link between information and problem solving activities of engineers. Beheshti (1989) implemented a log to uncover how much time was spent using library books. McKechnie (2000) used the diary-interview method as a secondary source of data collection. She asked mothers to keep diaries that reflect direct/indirect use of library materials by preschoolers and subsequently interviewed the mothers regarding diary entries. Mellon (1990) asked students to record their affective behaviours as well as details about the search process used. Kuhlthau (1993) acquired data on students perceptions of sources used/to be used from search logs. These previous studies have used log-style diaries or a composite of the intimate diary and memoir. In our research we did not want to impose the structure of logs, nor did we want the reflective commentary of a memoir; we wanted explicit details about use and thus, chose the diary-interview method.

STUDYING THE INFORMATION BEHAVIOUR OF ARCHIVAL USERS WITH DIARIES

The Study

In the first phase of our research, we want to answer the question: how do scholars use archival finding aids in the course of their research? Rather than narrowing our approach to a focussed study of finding aid use, we are starting our investigation at the point of entry into the environment using an approach that is typically found in systems design environments (Hackos & Redish, 1998). We are taking a holistic approach by first understanding users in their environment and the types of tasks that they attempt/accomplish within that milieu.

Research Design

Participants. We are recruiting ten history graduate students (primarily PhD candidates) to complete the diary component of our study. Graduate students spend a considerable amount of time in archives in the process of completing theses and are more likely to have a consistent use of archives over a specific period of time than other types of scholars. Scholars who already complete research diaries are unlikely to undertake the burden of an additional diary. Dairies completed by novice researchers when combined with the details from observations and the perspectives of interviewees will enhance our understanding of the process.

Procedure. We requested that the diary be taken to the archives for each visit and that 10-15 minutes be used during the visit to make an entry. We requested 10 instances of searching for something to be deemed 'completion' of the diary.

We provided participants with the following instructions:

> We would like you to make an entry each time you look for information in an archive. We are interested in anything that helped you to find a document. This could be a person (e.g., an archivist, another scholar), a catalogue, an inventory, a calendar, an index, or anything else that helps you to find information in an archive.

We requested that an entry could contain the following information:
- what tool (person, document, etc,) did you use to find information in the archives?
- how and why did you decide to use that tool?
- describe how you used the tool: what you did, what kind of information you found and what parts of the tool you used?
- what was helpful and what was not helpful?
- what other kinds of information would you like to see?

These items were selected based on the task analysis process used by Hackos and Reddish (1998).

We are making contact with diary-keepers three times during the process: one at the beginning which is primarily informational and instructional, one about halfway through for update and confirmation, and a third at the end for de-briefing. The form of the diary used by participants is a three-ring, one-inch-thick binder of 8½ inch x 11 inch pages. Each lined page prompts for date and institution visited.

IMPLEMENTATION

Recruitment, Retention and Finishing

In our first four months we recruited only one student and after a further six months had recruiting only an additional eight students. We find this very surprising. Our usual recruitment techniques from flyers placed in prominent places, notices posted to listservs, and personal contact with history departments and professors have been unsuccessful compared to the response received for other techniques that we have previously used, e.g., surveys, focus groups,

interviews and human experiments. Our initial request received only one response, which we subsequently used to pilot test our instruments. Since the completion of that pilot we have expanded our search through history departments on a national level, and to archives, and we have accepted as participants those using non-Canadian archives.

Among the set of eight participants, some who were recruited months ago, only one has submitted data. In addition to getting them to participate, is the difficulty in getting them to put pen to paper. Even though they agree that the task is worthwhile and the outcome of great potential to improving their future research experiences, there seems to be a reluctance to finish the job. Although technically we have not lost anyone, we have not collected much data.

Why have we had these problems? History graduate students at the point of actively conducting research tend to work solitarily with minimal contact with fellow students or indeed their professors. Gaining access to that community has proven to be cumbersome as the students are scattered around the world in a variety of archives and repositories. Secondly, students are reluctant to volunteer for a task that sounds onerous; our pilot for example took 6 months to complete and was terminated by us after 6 visits. Thirdly, we believe that the students are evaluating their own behaviour and determining what is important enough to record.

Length of Time to Keep Diary versus Unit of Observation

Previous diary implementations specified from one week (Zimmerman & Wieder, 1977; McKechnie, 2000) to three weeks (Elliott, 1997). A specific unit of time, e.g., hour, day, was inappropriate for our participants, as scholars may spend a week in the archives or not visit for a month or more. In addition, the number of trips made to an archive varies with the scholar's point in the research process, with the topic, and with the archives. Some trips may result in immersion in a box of records and documents (and thus of little interest to our research) while others may be limited to reading a finding aid.

In the pilot we arbitrarily specified 10 instances of searching for something to be 'completion' of the diary. But participants (to date) interpreted these instances as a physical visit. After the pilot we reduce the number of instances to 3.

The participants' misinterpretation of what constitutes an instance, however, has provided an added bonus. It is enabling data analysis based on visit. Each visit provides a holistic perspective on the user's experience in the environment and will enhance the data collected in interviews. More refined units such as an information-seeking incident or each event involving the use of an access tool will be identified within each visit.

Instructions, i.e., the Task Assigned

Our protocol based on task analysis (Hackos & Reddish, 1998) has met (to date) with mixed success for elucidating the process that we need to understand. In the interviews, we used the critical incident method but did not use that technique for diaries. A specific incident might not be a "critical incident" in the mind of a participant and a participant might not be able to assess the nature of the incident. Instead we aimed for inclusivity -- gather specific data on all types of

information gathering and use, and do so in an unstructured way. Many diaries used in previous studies exhibited log-like formats by providing specific categories in a form fill-in style of design. But we did not want to specify a defined linear or decision-tree like format with content categories, and thus influence in a major way the process that a novice researcher might take. Other than specifying the type of issues to include in an entry, we provided no direction on format.

Our pilot participant hand wrote 14 single-spaced pages covering six visits to five institutes over a period of five months. The details for initial visits were sparse reflecting the participant's uncertainty about what should be included as well as the novelty of the task; visits following the second interview were more comprehensive and seemed to change also with the participant's experience in using archives and confidence in that ability. This is not in keeping with the usual experiences that Verbugge (1980) found in diary keeping. Some entries were narratives based on the whole visit, somewhat in a memoir-like form; some entries were time-stamped much like traditional journal entries or daily schedule; and some entries were summaries of personal assessment/reflection on the visit. Much of the narrative particularly in the first few visits contained a chronological sequence of events at the archives: Who was consulted? What was consulted, recommended, used? This changed over the course of the visits as the participant injected personal summaries and recommendations for archives' tools, made comparisons between archives, and made the occasional wry comment. In contrast, our next participant delivered three single-spaced word-processed pages written weeks after the event, which were highly reflective and his interpretation of his experiences; this product was more similar to a memoir than a diary.

In hindsight we have questioned whether explicit categories would make the diary more efficient for entry and simpler for the participant. We are hoping for rich detail that describes at a micro level the experiences of an archival user and are reluctant to contaminate the process by imposing on participants our notions of how people interact with a finding aid. Anecdotally, our pilot participant reported that unstructured entries were simpler to do and enabled him to express thoughts and experiences. But clearly some form of training and more explicit instructions are needed to capture a rich process.

Use of Interviews

Regular interviews or follow-up help keep that participant on track and ensure that instructions are followed. But, more importantly, they add meaning to the diary entries and reduce/remove the potential for misinterpretation. Interviewees (to date) tend to be reflective of their archival experience. In the case of the second participant, he provided specific micro-level details on exactly how he used the finding aid, which was not present in his diary.

Format of the Diary

We initially explored the use of Palm Pilots and other digital gadgets for keeping diaries, before deciding on the paper-in-binder format. However we believe that this format works well for this group and confirmed by participants to date, although it is likely to be unacceptable to other

domain groups. As history students become more computer savvy, they will likely prefer to produce the diary in a digital form as one participant has done to date.

Alternative Approaches

Due to our implementation problems, we are also exploring the use of existing research diaries. In general all scholars maintain research diaries --- notes of their visits to archives and other repositories and their quest for data. Our initial request of a select number of archives was fruitless – they did not keep these logs or their authors did not deposit them. We managed to extract ten pages from one scholar who was reluctant to show us a research diary, but was willing to select a number of pages for us to examine. These pages were much like a log of a person's activities: which boxes were examined on a particular day, which finding aids were examined, which documents/copies were requested, to whom was letters sent, which appointments were made. Sometimes her frustration, and anxieties were present in personal reflections and sometimes there were innocuous comments about other people. But for the purposes of our study, if this sample is indeed a typical research diary, it lacks the level of specificity that we need. Instead of merely identifying which finding aid was used, we want to additionally know how it was used.

THE POTENTIAL OF DIARIES FOR STUDYING INFORMATION BEHAVIOUR

Based on our experiences to date, we believe that diaries have many pluses, but also many challenges to overcome in the study of information behaviour:
- A diary can be the voice and eyes of the participant – thoughts, perspectives – as both a participant and observer. Work within an archives and the multiple interactions with other scholars, archivists, finding aids, materials, inventories, etc. is a complex process that is not easily observable.
- The diary is a natural format; personal journals are an entrenched genre in our culture. The intimacy of a personal journal provides insights into thought processes and decision making because expressing that in a journal is a normal process. A diary entry is like a window into an event that is seemingly frozen in time. By maintaining the open-ended journal format, participants are encouraged to write from a personal and reflective perspective:
- However, despite the fact that a diary is being written, the participant still feels the pressure that the diary is in fact being written for a third party. How selective a participant will be in what he/she chooses to write versus what is left out cannot easily be determined.
- The diary encourages the inclusion of multiple types of content: specific details, problems, thoughts, analysis, and recommendations, at a detailed level of specificity and over time. In visit #6, the pilot participant said:
  > *I spent 10:30 through about 1:00 searching for Wood's documents at the online terminals, always using "Archivia-net". Archivia-net is a general online database which accesses almost all of what the [archives] has. It is divided into several "finding aids" which are sub-sections of the database based on general types of information (i.e., personal papers, government documents, etc.). Wood did not*

turn up when I searched the "General Inventory" aid using the terms H.W. Wood; Wood, H.W., Wood, Henry Wise, Henry Wise Wood, etc. (part of the problem an archivist told me was that this online system rejects periods (i.e., with initials) – a problem which doesn't exist on paper finding aids since one would do the narrowing down from a last name.

- The diary can encapsulate a lengthy, mostly non-observable process. Due to the nature of archives, the research process is iterative and unpredictable and may take an hour, a day, a month or sometimes years to complete. Participants can highlight in a detailed manner the significant events that factored into the search process as well as the events that were barriers to, or facilitators of, the successes. But training is essential to ensure that the level of detail is present.
- The diary is an uncontrolled and uncontrollable device. In interviews, focus groups and experiment, the researcher subtly by mere presence directs the proceedings. Diaries are completed remotely over a considerable period of time. Participants choose what to write, and when to write, and unfortunately choose what they perceive might be important enough to write. However, from those selective thoughts, we are able to better understand what participants perceive is important rather than just to respond to what researchers believe is important.
- Diaries as we propose to use them are a longitudinal device not confined to a specific week or month. Participants may lose interest over time, or lose focus on the task. Where is the boundary between rich information and scattered, mediocre details?

CONCLUSIONS

The value of diaries is in their ability to record events close to the time in which they occur from the perspective of the person involved. As such, they complement data collected by interviews, focus groups and observation. In our case, diaries have provided a window into a complex process that is difficult to observe. But the real handicap in capturing a timely process is in the commitment of participants to do the task. Diaries add another element to an already complex environment – from the nature of the research problem to the tools used, to the types of materials. In our research, diaries are but one technique meant to supplement and complement the types of data one collects in interviews and observation. To date it has proven to be a tedious but not unrewarding process to implement.

ACKNOWLEDGEMENTS

REFERENCES

Allen, T. J. (1966). *Managing the Flow of Scientific and Technological Information*. PhD. Dissertation, Massachusetts Institute of Technology.

Allport, G.W. (1943). *The Use of Personal Documents in Psychological Science*. New York: Social Science Research Council.

Beheshti, J. (1989). A longitudinal study of the use of library books by undergraduate students. *Information Processing and Management*, 25(6), 737-744.

Case, D. O. (1991). Conceptual organization and retrieval of text by historians: the role of memory and metaphor. *Journal of the American Society for Information Science, 42*(9), 657-668.

Cole, C. (1998). Information acquisition in History Ph.D. students: inferencing and the formation of knowledge structures. *Library Quarterly, 68*(1), 33-54

Conrath, D.W., Higgins, C.A., & McClean, R.J. (1983). A comparison of the reliability of questionnaire versus diary data. *Social Networks*, 5, 317.

Corti, L. (1993). Using diaries in social research. *Social Research Update*, Issue 2. Retrieved August 25, 2000; http://www.soc.surrey/ac.uk/sru/SRU2.html.

Davis, D.L. (1971). New approaches to studying library use. *Drexel Library Quarterly*, 7(1), 4-12.

Elliott, H. (1997). The use of diaries in sociological research on health experience. *Sociological Research Online*, 2(2); http://www.socresonline.org.uk/socresonline/2/2/7.html .

Fothergill, R.A. (1974). *Private Chronicles*. Oxford: Oxford University Press

Goodall, D. (1994). Use of diaries in library and information research. *Library and Information Research News*, 18(59), 17-21.

Hackos, J.T. & Redish, J.C. (1998). *User and Task Analysis for Interface Design*. New York: Wiley

Hilton, M.E. (1989). A comparison of a prospective diary and two summary recall techniques for recording alcohol consumption. *British Journal of Addiction*, 84, 1085-92.

Kuhlthau, C. C. (1993). *Seeking Meaning: A Process Approach to Library and Information Services*. Norwood, NJ: Ablex Pub. Corp.

McKechnie, E.F. (2000). Ethnographic observation of preschool children. *Library and Information Science Research*, 22(1), 61-76.

Mellon, C. (1990). Naturalistic Inquiry for Library Science. New York: Greenwood Press.

Plummer, K. (1983). *Documents of Life*. London: George Allen and Unwin.

Silberstein, A.R. & Scott, S. (1991). Expenditure diary surveys and their associated errors. In *Measurement Errors in Surveys* (pp. 303-326). New York: Wiley.

Sorokin, P.A. & Berger, C.Q. (1939). *Time-Budgets of Human Behavior*. Cambridge, MA: Harvard University Press

Verbrugge, L.M. (1980). Health diaries. *Medical Care*, 18(1), 73-95.

Wood, D.N. (1969). Discovering the user and his information needs. *ASLIB Proceedings*, 21(7), 262-270.

Zimmerman, D.H. & Wieder, D.L. (1977). The diary: diary-interview method. *Urban Life*, 5(4), 479-499.

Beyond Logs and Surveys:
In-Depth Measures of People's Web Use Skills [1]

Eszter Hargittai
Princeton University

Abstract

Finding information on the Web can be a much more complex process than previously possible on many pre-Web information retrieval systems given that finding content online does not have to happen via a search algorithm typed into a search field. Rather, the Web allows for a myriad of search strategies. Although there are numerous studies of Web search techniques, these studies often limit their focus to just one part of the search process and are not based on the behavior of the general user population nor do they include information about the users. To remedy these shortcomings, this project looks at how people find information online in the context of their other media use, their general Internet use patterns in addition to using information about their demographic background and social support networks. This paper describes the methodology in detail and suggests that a mix of survey instruments and in-person observations can yield the type of rich data set that is necessary in order to understand in depth the differences in people's Web search skills.

INTRODUCTION

There are many studies that look at how people use information retrieval systems and, in particular, how people search for information on the Web (for a review of this literature see Jansen and Pooch 2001). The Web Use Project at Princeton University adds to the literature on information retrieval in the following three important ways: 1. it considers the search patterns of users drawn from the general population instead of solely relying on people in the academic and information science communities for data; 2. it collects data not only on users' search activities but also on their use of other media for information retrieval, their demographics and their social support networks; 3. it recognizes that with the Web, searching for information is no longer limited to entering search queries in a search engine, rather, there are numerous ways in which one can go about finding information and these ways may lead to different results and differences in the efficacy of the particular information retrieval technique used.

This paper describes a method that allows us to measure differences in people's skills with respect to Web use and also presents findings from observation sessions and interviews with sixty-three participants. First, I briefly discuss the current state of research on Web use for information retrieval. Next, I discuss why the approaches in the existing studies are not always suitable to gaining a refined understanding of the differences in how people locate content online and how this project remedies some of these limitations of existing studies. Then, I describe the project methodology including a discussion of the sampling technique, technical specifications of the project, a detailed description of the two survey instruments, and the in-person observation

session of people's online search techniques. Finally, I present findings based on the first sixty-three people who participated in the study.

EXISTING RESEARCH ON WEB USE FOR INFORMATION RETRIEVAL

Scholars from many fields have explored how people use the World Wide Web for information retrieval. Advertising and marketing specialists often refer to users as "consumers" emphasizing their particular interest in people's online actions, namely their shopping behavior (Jarvenpaa, Sirrka and Todd 1996, Bell and Tang 1998). These studies often analyze users' behavior on only one particular site as opposed to exploring users' overall online behavior. Moreover, their sole interest is in how people decide to make online purchases, what influences these decisions, and how much shopping people engage in.

Much work conducted in the human-computer interaction field also tends to concentrate on particulars. Researchers in this area analyze people's use of specific design features and distinct Web site layout (see, for example, the Special Issue on World Wide Web Usability of the International Journal of Human-Computer Studies (1997)). Furthermore, they also look at features of software programs to assess important usability issues (see, for example, Greenberg and Cockburn (1999) for a detailed discussion of the "Back" button on browsers).

Alternatively, computer scientists draw on large-scale aggregate logs about people's Web use by analyzing all Web activity over a specified period (Catledge and Pitkow 1995, Huberman et al 1998.). An important limitation of many such studies is that they concentrate on the behavior of a small segment of the population by limiting participants to university faculty and students (e.g. Catledge and Pitkow 1995) or long-term users from the information technology profession (e.g. Choo, Detlor and Turnbull 1999). Admittedly, concentrating on such groups may be legitimate depending on the research questions; however, such sampling techniques limit the extent to which findings can be generalized to a larger segment of the Web user population. In cases where data are derived from larger segments of the online population (e.g. Huberman et al. 1998, Hoelscher 1998, Silverstein et al. 1999, Jansen, Spink and Saracevic 2000), there is no information about specific users and thus it is impossible to make any claims about how attributes of users may be related to their online behavior.

Private research corporations collect data on what sites people visit and how much time they spend on each page (e.g. CyberDialogue and MediaMetrix collect Web behavior information this way). However, such information is proprietary and does not include information on what users are actually looking for (if anything) and whether they are satisfied by the options presented to them on the screen. These data sets also do not contain information on how users perceive what they see and how they make the particular choices in their linking behavior and search strategies.

Researchers in the library and information science community have also conducted numerous studies on people's use of library resources that are often increasingly run on Web-based applications. Abramson looked at how people used the Web at public access computers by recording logs of use via a computer connected to the machines she was observing (1998). However, she only collected information about visited sites and time of day and week without

any information about users. Numerous case studies exist on the implementation of specific search programs in libraries (e.g. Payette and Rieger 1997) but these also limit their scope to a distinct user base and Web search protocol or library interface. There are also many studies (Hsieh-Yee 1993, Koenemann and Belkin 1996, Siegfried, Bates and Wilde 1993) that look at searches performed on various information retrieval systems (pre-Web applications as well), however, they focus on the details of query specifics (e.g. number of queries in the data set, session length, query length, use of advanced search functions) without considering information about user demographics or other information retrieval practices of the users.

Closest to the methods presented in this paper are some of the in-person user studies that have been conducted by library and information science researchers. Wang, Hawk and Tenopir (2000) collected data by observing how respondents search for information specified by the research team. Their project generated synchronized video-audio data, which were then analyzed for detailed information about respondents' search techniques. However, as often is the case in such studies, the participants for the study were graduate students and faculty in an information science program. In order to gain a better understanding for how the general population is using the Internet, it is important to include people from beyond the academic community in such studies.

The methods used in the studies cited here provide important information for a baseline understanding of how certain people navigate particular parts of the Web. As Friedlander (1997) points out in an editorial about research on digital libraries, referring to the 1970s debates about different methodologies; "quantification was sometimes a good thing and sometimes not; the value of the research resided in the questions it asked and the integrity of the findings, not the methods by which those findings were obtained." There is strength in both quantitative and qualitative approaches to people's use of computers and online resources. Existing studies either limit their scope to specific user populations (e.g. IT professionals or people who go to libraries), do not collect background information about user attributes, or look at use patterns on an aggregate level without collecting data about the specific goals of a Web session. The next section describes the methodology in detail followed by some preliminary data analyses.

COLLECTING IN-DEPTH DATA:
STRUCTURED OBSERVATIONS AND INTERVIEWS

Sampling

In order to be able to generalize from the findings, it is important to conduct the study on a random sample of users.   The Web Use Project looks at the online use patterns and skills of a randomly selected Internet user population. A random sample of residential addresses for Mercer County, New Jersey is obtained from Survey Sampling, Inc. and is checked against the National Change of Address Database maintained by the U.S. Postal Service. Potential respondents are first contacted via postal mail. They are sent a letter explaining the project and requesting participation with a brochure that presents more details about the study. People are also pointed to http://www.webuse.org on the Web for more information and are given the option of calling/writing to the research center to schedule an appointment. A few days after the letters

have been sent, the households are contacted by telephone. People who are identified as Web users are invited to participate in the study. The eligible adult (i.e. Internet user adult over 18) with the next nearest birthday is selected in order to randomly sample from within the household. [2] Web users are defined as people who go online at least once every month for more than using email. Although this is a low threshold for including people in the study, it is used to ensure that low frequency users who are nonetheless familiar with the Internet are also included. Asking questions about location of and motivation for use, low frequency of usage coupled with information about the user and his/her online behavior can shed light on some issues of inequality as these factors may influence one's skill sophistication in use of the medium.

People interested in participating are offered $40, which they receive after the interview session. Respondents are asked to come to the research site on the campus of Princeton University. [3] The respondents' email address is recorded and a time for the session is scheduled. Respondents are informed that they will receive a follow-up letter in the mail or an email message (based on their preference) with a reminder and directions to the research site. The day before the study, a reminder phone call is placed to the respondent or an email message is sent (based on their previously noted preference).

Technical Specifications

Two computers are used for the study to allow for variation in people's computer experiences. For PC users, the computer is a Pentium III running Windows ME connected to the university's network with a 17" monitor. For Macintosh users, there is a G3 iMac (15" monitor) also connected to the university network. The three most popular browsing software applications are all available on both machines (America Online, Internet Explorer and Netscape Communicator). [4] A relatively recent version is used although not a beta version to eliminate potential difficulties and ensure that users are not surprised by new features. Availability of both computer types and all three browser applications is important so respondents can use the program with which they are most familiar.

The computers are capable of connecting to the Internet on a T1 line. The HyperCam (Hyperionics 2001) software program is used to record the observation sessions on the PCs. This program creates audio-visual files (.avi) of the activity on the screen accompanied by the respondents' comments. HyperCam is able to capture action on the entire screen and thus serves as a very helpful tool for this study. A similar program, SnapZPro (Ambrosia Software 2001), is used on the iMac. Additionally, a program called Don't Panic (Panicware 2001) is used on the PCs to erase the browser and URL history on each browser program so that each respondent starts out with a clean slate and is not presented with auto-complete options if and when s/he types URLs directly in the location bar, pulls down the location bar menu to select an option or types terms in a form. [5]

The advantages of using such a recording program are manifold. The set up time and program costs are minimal. Although a microphone is used, otherwise the recording is unobtrusive (e.g. it does not require a video camera pointed at the respondent, rather, it records screen shots and verbal communication). The software runs on the machine that is being used so there are no interfering variables. The creation of .avi (audio-visual interleaved) files allows for easy access

to data. Storage is also straightforward and allows retrieval of data from numerous locations. Files are stored in a password-protected directory of the university network that is backed-up nightly offering safeguards against data loss. Additional backup copies are created on CDs. The biggest expense associated with this method (other than compensation to respondents) is the server space required for data storage. A session creates an approximately 500MB-1GB file which can be compressed to 200-600MB. [6] Although this is still a considerable amount of data if one does not have central server space allocation, nonetheless, in the age of read/write CD drives even such file sizes are manageable.

Interviews

The in-person sessions start with a twenty-minute interview about basic Web use questions. This interview draws on the Internet Module of the General Social Survey (GSS) 2000. The GSS is conducted every few years on a random sample of the American population with a response rate (70-80 percent) rarely achieved by other surveys. The GSS interviews are conducted face-to-face with people in their homes. For decades, the core section of the questionnaire has been replicated on every survey allowing for time-series analyses about people's political and religious beliefs in addition to a myriad of other attributes.   The GSS also contains topical modules that differ from year to year. In Spring 2000, a twelve minute Internet Module was added to ask people about their Internet use at home, work, school and other locations (e.g. libraries). [7] Questions were asked about what online services people use, what types of sites they visit, and how they use the Web for political and cultural activities.

The questionnaire presented to respondents in this study replicates sections of the GSS Internet module in order to allow for comparisons with a larger population of users. The questionnaire is administered verbally to establish a rapport between the researcher and the respondent. Administering the questionnaire right before the observation session proves to be very useful. Because the questions explore many facets of Web use, respondents are prompted to think about numerous details of their Web experiences before sitting down at the computer and embarking on the tasks presented by the researcher. For example, the questionnaire asks users to "think about the last time they viewed content online about a political or social-policy issue, current affair, or a political campaign". Then several questions are posed about this specific experience. By allowing users to reflect on such past experiences, they are given a "warm up" time before their Web use observation begins.

Details are also collected about the types of Web sites users visit in order to know with which types of sites they had had experience prior to the session as opposed to sites they are completely new to during the observation. This is important, as someone who often visits political sites is likely to exhibit different browsing strategies while searching for such Web sources, not necessarily because of a general higher level of skill in this task, but because of prior experience. Information is also collected on people's use of other communication technologies to offer a baseline comparison for how people use the Internet in their day-to-day lives. There are several questions that ask about people's social support networks both online and offline. This information is collected in order to understand how people's networks may influence people's ability to perform online tasks. Again, by asking these questions before the task session, people

are prompted to think about the various ways in which they find information including recommendations from friends, family and colleagues. [8]

After the observation session (described below), another questionnaire is administered collecting basic demographic information about the respondents (variables such as education, race, family income) in addition to data about the details of their Internet access. Information is asked about the quality of people's connections partly measured by their experiences instead of relying on their technical know-how regarding the modem or connectivity specifications. People's frequency of use is also assessed through several questions, as are their available social support networks. Core demographic information is collected to serve as independent variables in the analyses exploring the determinants of level of sophistication in Web use. Some of these latter questions are also drawn from the General Social Survey, some are drawn from the HomeNet study at Carnegie Mellon University, while others were added specifically for this study. A long list of computer and Internet related terms is also included on the survey and respondents are asked to rank their understanding of these terms. [9] The goal is to see whether the level of skill measured by analyzing people's actions online correlates with people's scores on these knowledge variables. Because the above method is time- and labor intensive and costly, a longer-term goal of this methodology is to suggest ways in which people's skills can be assessed via survey questionnaires instead of always relying on such detailed studies for this type of information.

Observation Sessions

The observations are conducted at a university research site. This approach has both advantages and shortcomings. Requesting users to come to a location will affect response rates. It also places people in a location with which they are not familiar and requires them to use a computer that is configured differently from the machine they usually use for browsing. This may influence the results, as certain settings (e.g. the default homepage and bookmarks) are not equivalent to their own. However, this approach controls for quality of Internet connection, and hardware/software differences. It also allows us to concentrate on Web use knowledge in a setting that is equally different and new for all. Moreover, using one computer allows the setup of particular software applications that are required for data recording as described earlier. No default page is set on browsers in order not to influence respondents' initial actions once online. The session is started off by the researcher asking the respondent to recall – if possible – the default homepage on the computer s/he uses the most. The respondent is also asked to comment on how much the browser used in the study in front of him/her resembles the one s/he uses most frequently. That is, the respondent is asked whether s/he has personalized anything on the browser and whether s/he has any bookmarks/favorites set.

Users are given a list of tasks to perform online in order to see how they would find certain information online (see the Appendix for the complete list). These resemble the question on the GSS (Q19) that asks how users would go about finding information about a political candidate. [10] However, instead of the hypothetical question asked on that survey, the researcher is able to watch users go through the process of finding a page and take detailed notes on what they do. This proves useful as strategies that people describe verbally often do not lead to the expected results once pursued online. A respondent may say "I would go to site x and look it up there"

but once s/he tries the method, the results are often not satisfactory to complete the task. It may be that the site s/he imagined having a certain type of content does not contain such content. An alternative reason for a failed attempt is that the respondent is unable to navigate a site's content. Yet a different reason for an unsuccessful suggestion is that the respondent is not even capable of arriving at the destination site (either because such a site does not exist or s/he cannot get to it). Thus, both respondent and researcher are able to see real time that the hypothesized response may not be an adequate one and the respondent can then pursue other avenues.

Data are also collected about duration on different pages, the amount of time it took to get to a site, and all actions taken to reach the destination. Detailed information is coded on each action including whether the person clicked on a link (and whether that link was an advertisement, an image link, a within page or within frame link), whether the person typed in a URL, what type of search engine the user pulled up, and all of the pages the user consulted before reaching the sought information.

Although aggregate level recordings of users' visits give us information about what pages users see (as per the type of data collected by commercial marketing corporations and some studies), they give us no information on what type of content users were actually looking for and how satisfied they were with what they found. Moreover, most logs of uses do not record information about moves that concern the local cache of the machine. Consequently, such data sets miss information about details such as the use of the "Back" button on browsers, which is a considerable problem given that its use comprises up to 30 percent of people's browsing activities (Tauscher and Greenberg 1997) and may be considered a part of one's level of search sophistication.

The task-oriented method is repeated for several types of sites such as arts, current events, volunteer organizations, shopping, music, health-related and job search services. This is important in order to gather information on universal versus topic-specific search strategies. Someone who is universally skilled in finding information may have highly sophisticated skills in locating any type of information whereas topic-specific skills imply that the user has considerably different search skill levels depending on the topic being sought. An example of topic-specific search skills is someone who possesses sophisticated methods for finding Web sites on online music, but has little knowledge of how to arrive at Web pages with scientific information about a health concern. Some of the tasks were familiar activities to some respondents but not to others. However, some of the tasks were new to all respondents which allows for comparison across cases with respect to a formerly uncharted territory (the two tasks that caught most people off guard were the one that required searching for a site that compares presidential candidates views on abortion, and the other that asked respondents to look for a page that displays "art by kids"; see questions six and fourteen in the Appendix).

During the session, the respondent is encouraged to make comments about his/her actions. S/he is asked if the actions s/he has been asked to perform are ones s/he has performed before. If the respondent gets enthusiastic about showing off a skill or search that s/he has recently engaged in, the researcher encourages her/him to do so even if this action is not directly related to the specific task at hand. Such actions add to the data about users' usual experiences and offer a baseline for what types of online activities users are already familiar with and are likely to

perform on a regular basis when not in a study setting. The researcher sometimes also asks clarifying questions right after the various search sessions. By asking questions about a search session right after it has been conducted and with an exact record of what had happened, it is possible to supplement browser destination data with a better understanding of what drives users' online actions. Any more general queries about the users' online behavior are left for discussion at the end of the session to make sure the exchange does not influence the respondent's actions during the latter parts of the study. By talking to people, we learn more about their actions and motivations than if we were simply observing recordings of the pages they visit. In other words, this project is not simply about studying people's sequence of use but also their search strategies, the underlying motivations of their actions, and their levels of satisfaction with their Web experiences. Information is also collected on what alternative ways – other than the Web – respondents may use to find the requested information. This, again, offers a baseline comparison for how much and for what purposes people tend to use the Internet in their everyday lives.

ANALYZING THE DATA

Tabulating Survey Responses

The above methods yield a wealth of data. Answers to the verbally administered questionnaires are recorded with pencil on paper and have to be entered into a database. The second survey is presented to the respondents online and thus the responses are automatically coded and can be downloaded in a spreadsheet format. [11] These data are then analyzed via quantitative methods. Each respondent has a corresponding user identification number that is used to match up the three components of the study. The data from the two questionnaires are merged into a data set. Once coded, the data from the search sessions are also added to this data set.

Classifying Search Strategies

The audio-visual (.avi) files generated by the screen capture application are coded while being viewed with a multi-media program such as QuickTime or Windows Media Player. The coding of search strategies is partly based on an exhaustive (albeit not mutually exclusive) list of ways in which one can arrive at a Web site (Hargittai 2000) summarized here:

Previous knowledge of a page:
1. Using the browser's default page.
2. Typing in the URL of a previously visited page, an address that one retained.
3. Using an entry from one's bookmark/favorites list.

No previous knowledge of a page:
1. Trying to guess the address of a Web page (e.g. by using the word of the sought information in the URL).
2. Doing an open search with the help of a search engine by typing certain terms into a search form (or in certain version of browsers and in the case of AOL, typing the term in the location bar itself, which comes with a built-in search engine).

3. Clicking on a directory category (or Web guide or AOL channel) and finding links via use of directories and subdirectories.
4. Clicking on an advertisement.
5. Recalling an address from exposure through another medium (e.g. radio, television, newspaper ad or article, billboard ad, recommendation from a friend or colleague)
6. Clicking on a link in an email from a friend/colleague/mailing list that contains a site recommendation.
7. Using browser add-ons for link (e.g. NeoPlanet or Alexa).
8. Clicking on a link from the contents of another page (other then options 3 & 4 above).

In addition to noting the number of the above strategies used by the respondent for the various searches, the researcher also takes note of other aspects regarding the respondents' online actions and comments about the search. Examples of these include the use of various browser features (the use of the Back button vs. the History file or clicking on links twice vs. once, pressing Enter vs. clicking on Submit and such details). Moreover, the researcher codes the route taken to the sites by making note of the various sites visited and the types of options that were chosen along the way (e.g. advertisement links, search results, etc.). A note is made about whether the respondent knows how to read URLs (Uniform Resource Locators or Web addresses) and knows how to interpret and manipulate them, whether the respondent knows who is responsible for the various Web sites s/he visits and which parts of the page the respondent looks at. We know such information by prompting the respondents to talk through their actions. Finally, information about search strategies is matched with how easily and quickly the respondent was able to find the desired information. Again, it is important to note that while someone can be extremely skilled at locating one type of information, they may have fewer skills in locating a different type of information. For this reason, both topic specific and universal search strategies are noted.

FINDINGS

The Sample

The findings reported here are based on 63 interviews conducted during Summer 2001 in Mercer County, New Jersey a mostly suburban community located about midway between New York City and Philadelphia with a total population of 350,000 living in approximately 126,000 housing units. All households have an equal chance of being included in the sample and people are chosen at random from within households. The response rate is approximately sixty percent.

Respondents range in age from 18-81 (four percent are in their teens, thirteen percent in their 20s, seven percent in their 30s, eighteen percent in their 40s, nine percent in their 50s, eight percent in their 60s, three percent in their 70s and one person in his 80s). Forty-three percent of this sample is male. Fifty-seven percent of these participants work full time and an additional eight percent work part-time. Their occupations are varied including real-estate agents, environmental policy analysts, blue-collar workers, office assistants, teachers, service employees and medical professionals in addition to students, unemployed and retired persons. Eighty-eight

percent of the respondents are White, six percent are African American and four percent are Asian American. Twenty-six percent live alone, 53 percent live with a spouse; the rest live with roommates or others (parents in most cases). The family income of these people is larger than the national average although it is important to note that Mercer County is one of the highest income counties in the country (moreover because this study excludes inner-city Trenton, the poorest neighborhoods are not in the sample at all).

Users' Skills

There is a large variance in the amount of time people take to complete all seventeen tasks ranging from twenty minutes to over 100 minutes. Most people are eventually successful in locating most of the requested content although some fail in succeeding with as many as half of the tasks. Respondents were encouraged to continue searching without giving up too easily (a minimum of five minutes was given for each task unless the respondent exhibited frustrations and expressed a need for moving on in which case the researcher read the next task).

Some patterns are emerging with respect to what makes a good searcher, which will be helpful in coming up with ways to educate users about their online actions. Although there are a myriad of ways in which one can improve one's sophistication in finding information online, there are a few basic skills that significantly improve the chances that one will find what one is looking for. Knowing when to use an open-ended search versus browsing through category directories compiled by large sites can help significantly. When looking for something more general such as comparison shopping for a used car, turning to categorized directories and Web guides offered by big sites may be more helpful than doing an open-ended search on a search engine. However, when looking for something very specific such as recipes for lactose-free cooking, a multi-term search will be much more efficient. Knowing some of the intricacies of how to use a search engine can be extremely valuable as well (e.g. use of Boolean operators). People who recognize the value of typing in more than one search term especially in the case of a complex search have a much easier time finding sites that address their queries. Moreover, understanding how search engines rank pages, and being able to read search results (including the URLs (Web addresses) of the results) can be quite valuable. When encountering a "Page Not Found" error, knowledge of URL truncation can be especially helpful.

An interesting finding of this study is the extent to which members of the general user population lack the basics of surfing the Web. A few people barely know what a Back button is and thus have an incredibly hard time moving from screen to screen. Many people know of no search engines and solely rely on functions of their browsers or Internet service providers. Some respondents also have a hard time entering valid search terms. One recurring mistake is entering multiple term queries without any spaces. When asked about this practice, one respondent replied that you are not supposed to use spaces on the Web thus the exclusion of spaces in between search terms. [12] Others' exhibit the exact opposite behavior by typing search terms in the location bar itself. However, given that most browsers now automatically redirect those terms into a search engine, this seems to cause fewer dead-end sessions. Yet others confuse a search bar with the location bar and type specific Web addresses in the search form. As several search engines do not seem to be configured to guess these addresses, users sometimes fail to reach their destination despite having had a specific URL in mind at the start of their session.

However, it is also possible that in a non-experimental environment (i.e. using the Web at home), these people may give up on their search due to frustrations. Several respondent at one point or another during the session offered the following type of comment: "now you see why I don't spend that much time online" referring to how frustrating search sessions can get and how hard it can be to find something.

In general, young people (late teens and twenties) have a much easier time getting around online than their older counterparts (whether people in their 30s or 60s). Some of this is clearly based on comfort with the technology they are using and not necessarily based on elaborate techniques they have mastered specifically with respect to the Web. Consequently, although older people (especially in their 60s and older) may be slower on the whole, age does not mean that they do not possess the necessary skills to be efficient searchers.

More findings are forthcoming in the talk.
Please check with author (papers@eszter.com) for latest version before citing!

APPENDIX – LIST OF TASKS PERFORMED DURING OBSERVATION SESSION

Participants are given the following instructions during the observation session:
> If you can recall, please bring up the page that is usually on your screen when you start using the Web. That is, the Web site that comes up when you start your Web browser program. Now please conduct the following tasks. Please note that there is no right way of doing these tasks. We are interested in seeing how you go about finding the following information online.

1. Where do you look for information about current events?
2. How would you find information about local cultural events in your area such as art shows, musical performances, theatre shows or movies? [wait for R to show preference among these tasks.] More specifically, I would like you to show me how you can find listings for local shows [movies, musical performances, theatre or art shows based on R's preference]. [13]
3. Where would you get sport scores online?
4. If you wanted to listen to some music online, where would you look for music that you could listen to right away?
5. Imagine you wanted to use the Web to find out information about a political candidate. What would you do?
6. Now let's assume that you wanted a neutral opinion comparing different presidential candidates' thoughts on a particular issue, say abortion. How would you find a neutral third-party site that compares presidential candidates' views on abortion? [If R challenges the existence of "neutral" sites on the Web then encourage R to focus on finding a site that compares the views of candidates even if the site is not necessarily neutral.]
7. Now think about health information. Let's assume that you or a friend/family member has been diagnosed with lactose intolerance. Where would you look for information about lactose intolerance?

8. As a follow-up, let's assume that you are having a friend over for a meal and this friend is lactose intolerant. You need to figure out what type of food you can serve. How would you find a recipe that is explicitly stated as acceptable for someone with lactose intolerance?

9. Now assume that you were interested in looking for a new job or exploring career opportunities. How would you use the Web for this purpose? [Prompt R to look for something realistic, a job for which they would be qualified/interested in. If R is retired and would not be looking for a job for themselves, encourage R to look for a job for a friend/relative.]

10. Let's assume that you have decided to buy a car and you would like to use the Web to help you in this. You are interested in getting a used car - say a 1995 Ford Escort - and would like to get some information about its features, its price, and most importantly whether there is one such car available for sale in your area. What would you do? [Be sure R looks for availability of such a car locally.]

11. If you were interested in finding the contact information (phone number, postal address, or possibly email address) of a long lost friend (a childhood friend or classmate), how would you go about it?

12. Let's assume that you have decided you would like to do some volunteer work. It is up to you what type of volunteer work you would do, or whether you would even have anything specific in mind. How would you use the Web to find information about volunteer opportunities? [Prompt R to look for something realistic that they might actually pursue. Pursue the interest until R reaches specific contact information.]

13. Let's assume that it is time to do your taxes. Where would you find the necessary forms and get information to help you with this process? [Ask R to look for the Federal 1040, the actual form.]

14. Imagine you would like to view some art by kids online. How would you go about this?

15. Have you ever viewed the pages of any museums or art galleries online? Can you show me where you would get such information?

16. Where would you find information about recall notices on children's toys? [wait..] Ideally, I would like to see a list of recalls, not just one.

17. Imagine that your wallet was stolen and you wanted to quickly find out what steps you need to take in order to avoid the potential problematic consequences of such theft. Where would you look for information about how to efficiently react to such a situation? In other words, you are looking for a resource that explains the three or four essential steps you have to take in such a situation, the phone numbers you have to call and such. Also ask: Is identity theft something you would be concerned about?

18. Please try to recall the last time you purchased something online. What was it? Where did you purchase it? How did you find that site?
   [This is not a task; it is for informational purposes about R's usual surfing activities.]

NOTES

their help with the survey components of the project and to Inna Barmash for her help with interviews and coding the data. Generous support from the Markle Foundation is kindly acknowledged. The project has also been supported in part by NSF grant #Z36240, a grant from the Russell Sage Foundation, and through a grant from the Pew Charitable Trusts to the Center for Arts and Cultural Policy Studies, Woodrow Wilson School, Princeton University. A similar version of this paper was presented at TPRC on October 29, 2001 in Alexandria, VA.

2. This study only includes adult English-speaking users. Two follow-up studies are already being planned for Spanish-speaking users and high school students.

3. Respondents are offered transportation if they cannot provide their own.

4. The research project has its own AOL account in order not to burden respondents' accounts with the time spent at the session.

5. A program with similar features was not found for the iMac and so the personal history is not cleared there. However, this influences few people as only two respondents out of 63 used the iMac.

6. Lower quality recordings may still be adequate for later analysis and will yield much smaller file sizes.

7. The mean response time of the Internet module was 12:26 and ranged from zero to 45 minutes.

8. See, for example, the following case: After being asked how she would find information about what to do if her wallet was stolen, a respondent offered the following information: "Oh see, I know all about that because on the Internet, a friend sent me a thing saying this is what to do if you ever lose your wallet. I've downloaded it and I keep it." It is important for respondents to signal these distinctions as it is not possible to replicate in the study setting users' email archives and their bookmarks/favorites listings yet it is important for the analysis of users' strategies to know that they do rely on recommendations from people they know for sites they visit.

9. A list of multiple-choice questions measuring people's actual knowledge of these computer and Internet related terms was added later in the study. The data reported in this paper include the responses to only three multiple-choice knowledge questions that were already administered in the first wave of the study.

10. The following is the specific wording of the question on the General Social Survey referred to here: "Imagine you wanted to use the Web to find out information about a candidate for political office. What would you do?".

11. The online survey uses the Princeton University Survey Facility, which is an application available to members of the Princeton University community for administering Web surveys (http://www.princeton.edu/~jkchu/Survey/).

12. As noted earlier, respondents were asked about some of their online actions after the full search session had been completed to make sure that the rest of their search behavior would not be influenced by the researcher's question.

13. Comments for the researcher are listed in brackets.

REFERENCES

Abramson, A.D. (1998). Monitoring and Evaluating Use of the World Wide Web in an Academic Library: An Exploratory Study. *Proceedings of the 61st Annual Meeting of the American Society for Information Science, Vol. 35*, 315-326.

Ambrosia Software. (2001). SnapZPro. http://www.ambrosiasw.com/

Bell, H., & Tang, N.K.H. (1998). The effectiveness of commercial Internet Web sites: a user's perspective. *Internet Research, 8*(3): 219-228.

Catledge, L.D, & Pitkow, J.E. (1995). Characterizing browsing strategies in the World-Wide Web. Paper presented at the Third International World Wide Web Conference (Darmstadt, Germany, April 10-14).

Choo, W.C, Detlor, B., & Turnbull, D. (2000, February 7). Information seeking on the Web: an itegrated model of browsing and searching. *First Monday, 5*(2). http://www.firstmonday.dk/issues/issue5_2/choo/index.html.

Friedlander, A. (1997, May). From the editor: design by experience. *D-Lib Magazine.* http://www.dlib.org/dlib/may97/05editorial.html.

Greenberg, S., & Cockburn, A. (1999). Getting back to back: alternate behaviors for a Web browser's back button. Proceedings of the 5th Annual Human Factors and the Web Conference (NIST, Gaithersburg, Maryland, USA, June 3, 1999). http://www.cpsc.ucalgary.ca/grouplab/papers/

Hargittai, E. (2000). Open portals or closed gates? Channeling content on the World Wide Web. *Poetics, 27*: 233-253.

Hoelscher, C. (1998). How internet experts search for information on the Web. Paper presented at the World Conference of the World Wide Web, Internet, and Intranet, Orlando, FL.

Hsieh-Yee, I. (1993). Effects of search experience and subject knowledge on the search tactics of novice and expeienced searchers. *Journal of the American Society for Information Science, 44*(3): 161-174.

Huberman, B.A., Pirolli, P.L., Pitkow, J.E., & Lukose, R.M. (1998). Strong regularities in World Wide Web surfing. *Science, 280*(5360): 94-97.

Hyperionics. 2000. HyperCam. http://www.hyperionics.com/

International Journal of Human-Computer Studies. (1997). Special Issue on World Wide Web Usability. http://www.hbuk.co.uk/ap/ijhcs/webusability

Jansen, B.J., & Pooch, U. (2001). A review of Web searching studies and a framework for future research. *Journal of the American Society for Information Science and Technology, 52*(3): 235-246.

Jansen, B., Spink, A., & Saracevic, T. (2000). Real life, real users, and real needs: a study and analysis of user queries on the Web. *Information Processing and Management, 36*(2): 207-227.

Jarvenpaa, S.L., & Todd, P.A. (1996). Consumer reactions to electronic shopping on the World Wide Web. *International Journal of Electronic Commerce, 1*(2): 59-88.

Koenemann, J., & Belkin, N. (1996). A case for interaction: a study of interactive information retrieval behavior and effectiveness. Paper presented at Conference on Human Factors in Computing Systems. Vancouver, Canada..

Panicware, Inc. (2001). Don't Panic 4.0. http://www.panicware.com.

Payette, S.D., & Rieger, O.Y. (1997, April). Z39.50: the user's perspective. *D-Lib Magazine.* http://www.dlib.org/dlib/april97/cornell/04payette.html.

Siegfried, S., Bates, M., & Wilde, D. (1993). A profile of end-user searching behavior by humanities scholars: the Getty Online Searching Project Report No.2. *Journal of the American Society for Information Science, 44*(5): 273-291.

Silverstein, C., Henzinger, M., Marais, H., & Moricz, M. (1999). Analysis of a very large Web Search engine query log. *SIGIR Forum, 33*(1): 6-12.

Tauscher, L., & Greenberg, S. (1997). How people revisit Web pages: empirical findings and implications for the design of history systems. *International Journal of Human Computer Studies, 47*(1): 97-138.

Wang, P., Hawk, W.B., & Tenopir, C. (2000). Users' interactions with World Wide Web resources: an exploratory study using a holistic approach. *Information Processing & Management, 36*(2): 229-251.

Complementary User-Centered Methodologies for Information Seeking and Use: System's Design in the Biological Information Browsing Environment (BIBE)

P. Bryan Heidorn, Bharat Mehra, Mary Lokhaiser
Graduate School of Library and Information Science, University of Illinois

Abstract

Complementary socially grounded user-centered methodologies are being used to design new information systems to support *biodiversity informatics*. Each of the methods--interviews, focus groups, field observations, immersion and lab testing--have their own strengths and weaknesses. Methods vary on their ability to reveal the automatic processes of experts (that need to be learned by novices), data richness and their ability to help interpret complex information needs and processes. When applied in concert, the methods provide a much clearer picture of the use of information while performing a real life information mediated task. This picture will be used to help inform the design of a new information system, Biological Information Browsing Environment (BIBE). The groups being studied are high school students, teachers and volunteer adult groups performing biodiversity surveys. In this task the people must identify and record information about many species of flora and fauna. Most of the information tools they use for training and during the survey are designed to facilitate the difficult species identification task.

THE CONTEXT

In order to understand the use of information it is important to understand the context of use; the task, the goals and skills of the participants. In this case we are studying people conducting biodiversity surveys. The primary goal of the people conducting the survey is to define a study area and then to document the organisms that live in that area. Students and adult volunteers need to perform many tasks as part of EcoWatch but one of the most challenging tasks is species identification. EcoWatch is a program run through the Illinois Natural History Survey. In this program, citizen volunteers, who are often not scientists, gather information about the biodiversity and health of natural areas. Participants need to be able to identify hundreds of plant and insect species. In this project, we are studying how the participants learn to identify species and when they fail. The objective of the BIBE project is to understand the information use and information needs of several groups of participants (Heidorn, 2001). The study of information use in this context is called *biodiversity informatics*. Biodiversity informatics is a broad area of research and application. The mission of biodiversity informatics is to provide information-related support to the process of identifying and quantifying biological diversity. The process of information gathering and analysis, as well as the information itself is used to support research, education and policy development. This information is critical to the understanding ecosystems and the impact of human and natural phenomenon upon those systems.

Many scientists believe that the earth is facing the loss of large numbers of species (Ehrlich and Wilson, 1991). For example, during the next fifty years up to two-thirds of plant and animal species, mostly in the tropics, may be lost (Raven, 1999). This extinction would have an impact on evolution that would last for many centuries (Myers & Knoll, 2001). This information is essential for gauging the planet's health yet little is known about ecology and biodiversity compared to user biological disciplines. "... Ecology and conservation biology are still disadvantaged. Their growth is hampered by a seldom-acknowledged deficiency: out of ignorance of the world's biodiversity, particularly at the level of individual species, where knowledge is fundamental to all other studies of diversity and hence of the whole living environment (Wilson, 2000)."

THE STUDIES APPROACH

We are closely studying a small sample of people involved in biodiversity surveys. To date this has included interviews with three teachers and three professional botanists. There have been three ForestWatch field observation trips videotaping thirteen students and two teachers. There have been two focus groups including a total of seven high school students. Two of the research team members were trained in ForestWatch and one in PrairieWatch. Both can be considered informants in the process. This represents dozens of hours of audio, video and field notes and hundreds of pages of transcripts but still is a small sample of the participants in the program. 340 Illinois high school teachers from over 300 different schools have been trained in at least one ecosystem monitoring program (RiverWatch, ForestWatch, PrairieWatch, WetlandWatch, and UrbanWatch) through the PLAN-IT EARTH partnership. As numbers are cumulative, we could conservatively estimate the EcoWatch program has reached between 15,000 and 20,000 students in classes taught by these teachers. In addition, hundreds of adult volunteers monitor hundreds of forest and river sites.

SOCIALLY GROUNDED USER-CENTERED METHODOLOGIES

It is obvious that going digital has as much to do with social transformation as technological conversion (Levy and Marshall, 1995). For any aspect of technology use is very much 'embedded' in real social settings and contexts (Nardi & O'Day, 1999). Thus, the design of technological systems for biological information and online visual browsing tools for biological information need to have the following characteristics:

- **They should be socially grounded and participative**: This is dependent on the fact that the success of a socio-technological system is based on the adoption by its users (Fischer, 1992).

- **They should be shaped by users' needs, expectations, and situations of use:** This is because these factors are significant determinants of the usability phenomena (Suchman, 1987).

- **They should accommodate social practices of users associated with system use:** The documentation of social practices should inform system design for successful implementation (Bowker, et al., 1997).

Multiple socially grounded, user-centered methodologies are being employed in the BIBE. These complementary methods often reveal detailed conditions of information seeking behaviors of users to achieve specific goals. These methods identify aspects about user dynamics, social roles, barriers faced, expectations of participation, role of prior training, use of resources, and impact of computers. No one method was able to provide all this information about the participants.

Each methodology has strengths and weaknesses for eliciting different kinds of information. For example, focus groups are considered particularly suitable for obtaining several perspectives about the same topic, though problems arise using this socially grounded research strategy when the researcher has to identify the individual view from the group view (Gibbs, 1997). In order to compensate for the lack of voicing in individual perceptions on some topics in the focus groups, a number of formal and informal interviews were conducted to hear concerns of various stakeholders. The following sections present contextual use and descriptions of some of the user-centered techniques that were employed in the BIBE Project.

Interviews, Botanists

There were two types of interviews conducted in this project. The lead questions for one set of interviews were designed for professional botanists. Six interviews were conducted at the Missouri Botanical Garden and the Illinois Natural History Survey in the summer of 2000. There was a fixed set of questions about sources of information for identification. They were ordered from the open-ended to specific questions. These included the design of identification keys, the selection of characteristics and character values for identification and the design of flora. Each interview lasted about two hours. Field notes were kept on the interviews. These interviews revealed how the experts who author paper-based keys intend for them to be used as well as some of their known shortcomings. The results of these interviews were used to design prototypes of identification systems that were used in the focus groups and interviews with teachers in the fall of 2000 and spring of 2001. The strength of these interviews was that they provided design input from people who have many years of experience with these information tools. One weakness is that the interviews provide only second-hand information about the end-users of the tools. Another weakness is that experts after years of effort have become so highly trained that many of the difficult processes involved in identification may have become automatic (Newell & Rosenbloom, 1981). This can make it difficult for experts to predict what will be difficult for a novice. This problem was reduced in our study since some of our interviewees were also teachers and are frequently reminded of these difficulties.

One interesting observation about the use of paper keys was that they require several minutes per species for an accurate identification. As was observed several months later during field observations, this proved to be a barrier to the use of the keys.

Interviews, Teachers

The second set of interviews was with high school teachers immediately after they had completed biodiversity surveys with their students. These interviews were formal and directed by a fixed set of questions that paralleled the questions used in the focus groups with students discussed in detail below. These interviews were recorded and transcribed. These interviews had the strength of providing insight into the unstated goals of project in the eyes of the teachers.

As might be expected, the goal was to teach students about ecology but more generally about the scientific method. Many additional classroom activities are tied to the EcoWatch fieldwork. Teachers used the PLAN-IT curriculum to insure that classroom activities met the Illinois State Teaching Standards (Lisowski, 1997). PLAN-IT was developed concurrently with ForestWatch procedures. Through the curriculum students learn concepts such as sampling technique, biomes and biodiversity. Some of these goals were apparent in classroom observations, some were not.

Another strength of the interview technique is that it allows the researchers to discover why the classes were organized the way they were. The "why" of activities is not apparent from field observation alone. Teachers are conscientious about the quality of the data being submitted to the IDNR. Some decide not to submit the data but to keep the survey as a purely academic exercise. Teachers were also concerned about the time constraints imposed by the classroom setting. After accounting for travel time, students only have 40 minutes to conduct the tree identification portion of a survey. Teachers compensate for this through classroom training, practice field trips to non-official ForestWatch sites and by assigning different survey activities such as site boundary layout, to different classroom sessions. In large classes teachers can assign one of the three 50 meter transects to a different group of students. The weakness of these interviews is that the results were mediated by the teacher's goals and desired outcomes. While the teachers were candid and honest in the interviews, because of the necessary time constraints of a one-time interview, the reports were oversimplified representations of the reality of the work being conducted by the students. This could be addressed through interviews with multiple teachers and through multiple interviews with the same teacher during different phases of the classroom activities.

Immersion

One method to gain a better understanding of user needs is to become a user (Beyer & Holtzblatt, 1995). In this approach system developers endeavor to have the same training, experience, and perform the same task as the people who will use the system. In immersion, rather than simply observing others performing the survey tasks, the designers perform the tasks themselves. ForestWatch volunteers are given a day of training on the survey procedures and species identification. After classroom work the volunteers go to the fields to conduct a practice survey. Bryan Heidorn, the Principal Investigator on the project, received ForestWatch training and became a certified ForestWatch volunteer in the spring of 2000. Likewise, Mary Lokhaiser, a doctoral student on the project, received Forest Watch and PrairieWatch training in the summer of 2000. Observing volunteers making mistakes in species identification and using information aids is very different from needing to distinguish between an American elm from a slippery elm yourself. These experiences serve as a first-hand source of understanding of the information needs of participants, including not only plant and insect identification information but also good sources of treatments for poison ivy, insect bites and sunburn.

A weakness of this approach is that it is often impossible to apply in real life situations because of costs, time or physical limitations. A white middle-class male designer can not become a poor, minority mother with an information need. (This it is wise to invite the users to participate in the design.) Fortunately, in the case of ForestWatch, it is relatively easy to become a volunteer. Another potential weakness of immersion is automaticity. As processes become automatic, as you become good at a task, you no longer recognize what was difficult in the beginning. Since training for all volunteers is brief this did not prove to be a real problem.

Field Observation

Field observation trips in the fall of 2000 and spring of 2001 provide examples of the techniques used in the study and the information we can gain by observing information use in context. The goal of the observations is to determine where the participants make mistakes, why the mistakes are made and find solutions. The field observations also provide information on volunteer interactions. The participants' discussion among themselves and their actions tell us what they consider to be important and what characteristics of the species they focus on.

The main goal of the participants was to fill out a data form that is used by the Illinois Department of Natural Resources. An example of part of one of the forms is included in Figure 1. The information they must gather includes the identification and size of all trees in each 10-meter section of three 50-meter survey areas. In the fall 2000 field observation, two researchers video-taped five Illinois high school class during a fall ForestWatch survey on one 50-meter interval while a third observer took field notes. The field notes included time stamps, and a sequential list of the species assigned to trees along with the speaker. The same team then recorded the behavior of two professionals surveying the same site. This observation served as a standard against which the volunteer participant judgments could be based. Two cameras are necessary because the participants, while working together, can be over 10 feet apart. Even with the two cameras it was sometimes difficult to determine which individual was speaking and how decisions were being made. The team was unable to match, tree for tree, the identification of the trees. To complete the analysis the research team had to return to the forest with another tree identification expert and create a map of the location of each tree and its species name. It was only then that we were able to complete analyses of the videotapes.

In the fall of 2001 we changes the nature of the field notes and as a consequence were able to reduce the number of cameras to two. The new field notes take the form of a rough map. While the volunteers work, the researcher plots the location of each tree along with the time and its volunteer assigned identification and size. During data analysis we were able to much more easily follow the video.

Figure 1: ForestWatch Data Input Form

47

The information resources used in the field include: laminated color pages with plant images and descriptions as in Figure 2; web descriptions of the tree which they had printed out as in Figure 3. These are provided by the IDNR. They also used field guides such as the Forest Trees of Illinois (Mohlenbrock, 1991).
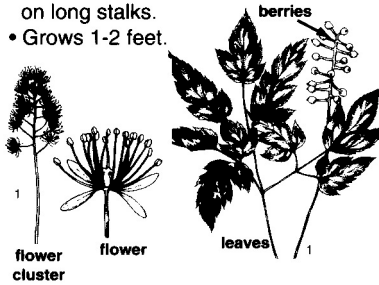
Figure 2: Laminated Sheet Plant Image

Figure 3: Web Resource

The students and teacher identified a majority of the trees correctly; however, researchers were interested in the process of identifying trees incorrectly in order to devise a system to improve EcoWatch data gathering accuracy. We were able to identify several error types that might be easily corrected.

The survey took the students 35 minutes to complete. In that time they identified 55 trees at a surprising rate of one and a half trees a minute. This conflicted with the professional estimates of how long it should take from the interviews but the teaches had trained the students in fast identification prior to entering the field. Only 42 trees were considered to be in the survey area. The remainder were outside of the area and should not have been counted. The identity of one of these could not be confirmed and there was one tree that may have been missed by the observers but the difference was mostly attributable to trees on the edges of the survey area. This was an unexpected discrepancy. This may indicate a need to change the techniques for laying out survey areas or changes in instructions to the volunteers. All further discussion in this paper is limited to the 41 trees that were identified by both the volunteers and the professionals. The professionals and the volunteers agreed on 33 of the 41 trees (80%). The following table is a breakdown of the differences between professional IDs and volunteer IDs.

Table 1: Misidentifications

Professionals	Volunteers
American Elm (3 instances of misidentification)	Slippery Elm (3 instances)
Catalpa	Sassafras
Sassafras	Ash
Dogwood	Sour Gum
White ash	Hickory
Flowering Dogwood	Alternate Dogwood

Of the eight errors, three were confusion of two elm species. It is relatively difficult to tell the difference between the American elm and the Slippery elm since the main differences are that the upper surfaces of slippery elm leaves are rougher than the American elm. From the videos we know that the professionals always felt the leaves or broke off a small sample of bark to see if it had the red line characteristic of the American elm. The volunteers never did this.

Some important design consequences from the field observations are that the student volunteers rarely consulted the information resources that they had on hand. Printouts of web-based tree IDs in Figure 3 were used twice. They identified almost all trees from memory. At least for trees, where it is possible to memorize species, it might be better to design systems that aid learning rather than systems to be used in the field. Students were impatient with the amount of time required to use keys and field guides. Any tool used in this context would need to be able to identify species in seconds rather than minutes. Another observation is that students follow specific social roles that are in part dictated by the mechanics of the situation. One person records data, another measures tree width, one reads coding information and one or two individuals are recognized as "experts" or "authorities" in identification. These mechanics and roles are important to consider in the new design.

The students are, however, incorrect part of the time. The focus groups conducted after the survey helped uncover some of the reasons for these errors. One reason for these failures may be overconfidence in their own knowledge and the knowledge of the "authorities."

The strength of field observations is that they provided an unfiltered view of the information task. Unlike the other methods used in this study, the field observations were not interpreted through the perception of the participants. Most user actions and all conversation were captured. The difficulty with the approach is the volume of the data. It is difficult to associate, one-for-one the opinions of the volunteers and those of the experts.

Focus groups

Powell & Single (1996) define a focus group as "a group of individuals selected and assembled by researchers to discuss and comment on, from personal experience, the topic that is the subject of the research" (pp. 499). Characteristic features of focus groups that proved significant in their use in the BIBE Project included organized discussion (Kitzinger 1994), collective activity (Powell et al 1996), social events (Goss & Leinbach 1996) and interaction (Kitzinger 1995). Focus groups in the BIBE Project were formal discussions organized around the biodiversity surveys in which the students had participated and interacted with each other. Relationship between the participants in the field and during the focus group discussions clearly expressed vibrant social dynamics in terms of role demarcations, designation of authority and communication and inter-personal exchanges.

Focus groups were conducted with students during Fall 2000 and Spring 2001 monitoring sessions. Additional focus groups will be conducted over the next two years. The focus groups are composed of two parts. The first part includes a series of questions about the classroom and ForestWatch survey activities. Significant issues addressed in the focus groups were: visions and expectations of participants; prior preparation and training activities; past experiences and procedures followed; distribution of roles and social dynamics; use of resources; barriers and what worked well in different situations; suggestions for improvements; and use of computers. The second part begins with a demonstration of an interactive key program, IntKey (Lawrence et al., 1999). The participants are then invited to use the program and are asked about their general impressions. This is followed by a series of specific questions about each of the major features of the program. Finally, having been primed as to some computer-based possibilities, the volunteers are asked, "What could be done to improve the survey process?" The focus groups also provided feedback about the use of a number of identification tools such as keys, the EcoWatch training manual, colored pictures provided in the EcoWatch training packet, other books and brochures, online sources, etc. that are employed during the biodiversity surveys conducted in EcoWatch programs. The focus groups also included having participants interact with computer-based keys. Transcriptions of audio recording provided direct quotations, detailed statements, key suggestions, and noteworthy experiences. Additionally, notes taken by researchers helped record the complexities of participants' experiences.

During the Fall 2000 monitoring survey, researchers conducted a one-hour focus group at a high school. Four female high school students of the twenty-plus students, who had participated in prior field observations, participated in the focus group. Researchers conducted an approximately one-and-half hour focus group during the Spring 2001 monitoring season at a high school in Central Illinois. Three male high school students participated in the focus group and narrated their experiences in past and present EcoWatch programs. These three students were the only students in the class that that conducted the survey. There was no indication that

the gender of volunteers had any impact on the focus groups. The following were some significant findings during these focus group discussions:

- Training before entering the field is essential for the task of plant identification. Students, including those with no prior experience, did learn a great deal about trees in the classroom and around the school before going to the field. One participant stated, "I have learned a ton from trees cause when I first came in I knew, I didn't know any trees at all and now I mean I know the main ones. You know, if it's a maple or it's an oak or I don't know you know specific types, but before I'd look at a tree and be like 'Oh, it's pretty. Don't know what it is, but...'"

- Students believed that computers, particularly resources available on the Internet, could have an important role in tree identification. As stated by one of the participants,
  "Well, also you could do it where you could... take that and like hook it up, like have the Internet going, and give you sites to where you can go and ask somebody, maybe like, or it has like helpful hints... or you can just like get on the Internet and like show them the leaves you know with the scanner or just compare."

- Traditional tree identification guides and how they were not as useful or as trusted as expert opinion, good quality photographs, and leaf collections. As one student stated, "It's easier for me to see it... instead of having someone look it up." Another student spoke of the significance of photographs and images as compared to traditional textual descriptions, "I mean that's the easiest way for me to see a picture, see somebody doing it, not written down."

- From notes made during the focus group discussion, some of the reasons why the students did not use the materials in the field included the following: students' over confidence in their knowledge as applied to plant identification; low motivation factors; the lack of appropriate, easy-to-use keys with color images; and time limitations for the tasks at hand.

- Volunteers quickly arrived at a consensus even when the consensus was wrong. Volunteers with a correct hypothesis for the ID of a species would too quickly yield to other opinions in the group without referring to the available paper resources to verify the ID. One solution might be to restructure both the social dynamics and the forms to help foster debate and verification. For example, at least two individuals might independently ID each tree. Software could point out discrepancies and species characteristics that would resolve the conflict.

From both focus groups conducted during Fall 2000 and Spring 2001 monitoring seasons, there were some important differences noted about the nature of students participating in the biodiversity surveys and the procedures adopted during those times. Briefly presented are two significant differences:

- Three students participated in the biodiversity survey conducted during the field observation in spring 2001. The same three participated in the focus group. Out of the three, two of the students had participated for the past three years in biodiversity surveys at the same sites. On the other hand, during the fall 2000 monitoring season, there were 27 students who

51

participated in this biodiversity survey. Of the five included in the focus group and field observation, none had previously conducted a survey.

- The nature of class settings of students participating in a biodiversity survey has a tremendous impact upon the social and group dynamics expressed in the field, the level of participation of the teacher in the plant identification process, and the demarcation of student roles while conducting the surveys. During Fall 2000 monitoring season, since there were many students on the site, it was practical to divide them into groups of 5-6 students each. The teacher assigned each group to one of the three transect lines. During the Spring 2001 monitoring season, since there were only three students participating, one group completed all the data collection along the three transect lines. The teacher was able to give more focused attention to the tasks being performed. In this situation, the teacher immersed herself in the process of identification and contributed actively in the process. She took the role of an expert or supervisor for the student group. Her decision was equally respected and rarely contradicted by group members. The teacher assigned one of the students, who was participating in the biodiversity survey for the first time, the role of recording. Since the other two students had participated in biodiversity surveys of the same site for the past consequent three years, the teacher had complete faith in their identification skills. In Fall 2000 monitoring season, the teacher was more involved in assigning groups and insuring that the groups completed the tasks at hand, and only occasionally became involved in the details of identification.   Consequently, each group had to assign different roles to its members and one of them took on the role of the expert or leader in the task of identification. Based on past experiences, this person's decision was often accepted by other students unequivocally, though there were some occasions where a discussion ensued before this person's decision was finally accepted. The other group members took on roles of recorder, verifier, measurer of plant characteristics, person involved in discussion, person carrying materials, etc. Sometimes the expert/leader and recorder or note-taker was the same person.

DIFFERENCES IN GOALS

There are sometimes conflicts between the goals of the main participants in EcoWatch. For examples, the IDNR scientists and the State legislature are interested in long term trends in the state environment. Consequently, they are not interested in looking at data from a single season alone but need information for five, ten or more years before it is useful. Adult volunteers, teachers, and students however frequently have more immediate goals. These groups are interested in the "their" plot of land or their stretch of river. They are concerned when an invasive species was identified. For example, on finding what they believe to be an invasive garlic mustard (*Alliaria petiolata*), they are tempted to pull it out (although we have not observed them actually doing so in their survey plots). These groups are also much more interested in seeing the fruits of their labors in the short term. All three teachers that were interviewed reported that they would like to see how their data is being used shortly after it is submitted.

CONCLUSION

Multiple user-centered methodologies are being applied in an ongoing project to design an information system to support biodiversity survey work by adult and high school volunteers. No

one approach or user group can provide all of the social and technical information needed to understand the user needs but the combination of approaches provide new insight into the use of information and the information needs of the users in this context.

There are many axes along which the effectiveness of methods for studying information seeking and use can be gauged. There are three main axes that are discussed here. These include controlled vs. automatic processing, interpretation vs. direct perception, data richness vs. interpretability. The user groups in the study varied in terms of the level of controlled vs. automatic processing and therefore the transparency of their responses. When people are initially learning a skill they must give conscious attention to the task in order to accomplish it. During this phase people can easily verbalize what they are trying to do since the tasks are frequently verbally mediated. Students, teachers and some adult volunteers fall into this group. This group is also the main group to be served by the new information services planned in this research. Focus groups, interviews and field observation can all reveal information about the information constraints on the task. Immersion of the designers into the information task forces the designers into a mostly controlled mode form of processing. This allows for more direct observation of the difficulties experiences by other novices. Unfortunately, the novice volunteers and the researcher/designers do not always know the "best practices" and sometimes cannot successfully accomplish the tasks. Experts such as botanists in interviews and field have the required knowledge but it can be tacit and automatic. Experts may not reveal critical knowledge during interviews unless directly prompted for information. For example, in the field experts frequently "just know" the identification of tree species. They may say, "That looks-like a sugar maple." without verbalizing what it is that makes it look like a sugar maple. Field observation reveals some of this knowledge. For example, when the experts are identifying elms they peel off pieces of the bark. The action indicates a critical piece of information, orange and white, stripped bark indicates an American elm. Addition insight can come from experts through the complementary, novice interviews and field observation. Questions and errors from the novices and be answered by the experts.

The next major axis is interpreted vs. direct perception. Some methods are good at eliciting information about frequently unspoken goals and motives. Interviews with scientists and teachers provide this sort of information. However this level of interpretation is also potentially subject to reporter bias. The interview can lead to oversimplification or an idealized view of a complex situation. This method is complemented by field observation of novices that can expose these oversimplifications through direct observation. At the same time the results of professional interviews can form a framework for interpreting the immense amount of data that can be collected in field observations. This dimension is correlated with axis of data richness vs. interpretability.

Field observation can provide a rich source of data for design. In our case it included field notes and maps as well as video. We were able to identify the information resources used in the task including field guides, ForestWatch identification keys and reports of prior surveys. The difficulty comes in interpreting the data and deciding how this interpretation should affect design. Focus groups with the novice volunteers after surveys complements this data, helping to determine what is important and what is not, what is difficult and what is easy for novices. Immersive design allows observes to understand better the non-obvious mental and emotional

processes of the volunteers. Separate interviews with experts help to fill in missing knowledge of the novices.

Interviews, focus groups, immersion and field observations of different groups of users give complementary perspectives on the information needs of people performing information dependent tasks. Each method can be used to compensate for some of the weaknesses of the other methods.

ACKNOWLEDGEMENT

REFERENCES

Bowker, G., Star, S., Turner, W., & Gasser, L. (editors). (1997). *Social science, technical systems and cooperative work: Beyond the great divide.* Hillsdale, N.J.: Erlbaum.

Beyer, H. R.& Holtzblatt, K. 1995. Apprenticing with the customer. *Communications of the ACM*, 38(5), 45-52.

Ehrlich, P.R. & Wilson, E.O. 1991. Biodiversity studies: Science and policy. *Science* 253:758-762.

Fischer, C. (1992). *America Calling: A Social History of the Telephone to 1940.* Berkeley, CA: University of California Press.

Gibbs, A. (1997). Focus Groups. *Social Research Update*, Issue nineteen, Winter 1997.

Goss, J., & Leinbach, T. (1996). Focus groups as alternative research practice. *Area*, 28 (2): 115-23.

Heidorn, P. Bryan. (2001) A Tool for Multipurpose Use of Online Flora and Fauna: The Biological Information Browsing Environment (BIBE), *First Monday*, 6(2) (February 2001). [http://firstmonday.org/].

Kitzinger J. (1994) The methodology of focus groups: the importance of interaction between research participants. *Sociology of Health,* 16 (1): 103-21.

Kitzinger J. (1995). Introducing focus groups. *British Medical Journal*, 311: 299-302.

Lawrence, J. F., Hastings, A. M., Dallwitz, M. J., Paine, T. A. and Zurcher, E. J. (1999). 'Beetle Larvae of the World: Descriptions, Illustrations, Identification, and Information Retrieval for Families and Sub-families.' CD-ROM, Version 1.1 for MS-Windows. (CSIRO Publishing: Melbourne.)

Lisowski, M. (1997) PLAN-IT EARTH Takes Teachers Out of the Classroom & Into the Environment. *Spectrum,* 23(4): 35.

Mohlenbrock, Robert. *Forest Trees of Illinois*, Eighth Edition,. Illinois Department of Natural Resourses, Division of Forest Resources and Natural Heritage.

Myers, N. and Knoll, A. (2001). The biotic crisis and the future of evolution. *Proc. Natl. Acad. Sci. USA*, Vol. 98, Issue 10, 5389-5392, May 8, 2001 [http://www.pnas.org/cgi/doi/10.1073/pnas.091092498 (Accessed June 1, 2001)].

Nardi, B. & O'Day, V. (1999). *Information Ecologies: Using Technology with Heart*. MIT Press.

Powell, R., & Single, H. (1996). Focus groups. *International Journal of Quality in Health Care*, 8 (5): 499-504.

Powell, R., Single, H., & Lloyd, K. (1996). Focus groups in mental health research: enhancing the validity of user and provider questionnaires. *International Journal of Social Psychology*, 42 (3): 193-206.

Raven, P. H. (1999) Plants in peril: What should we do*? 16th International Botanical Congress*, Missouri Botanical Garden, St. Louis, MO.

Suchman, L. (1987). *Plans and Situated Actions: The problems of human-machine communication.* Cambridge: Cambridge University Press.

Wilson, Edward O. (2000). A global biodiversity map. *Science*, 289: 2279.

Following Experts at Work in their Own Information Spaces: Using Observational Methods to Develop Tools for the Digital Library

Paul Gorman, Mary Lavelle, Lois Delcambre, David Maier
Oregon Health and Science University
and
Oregon Graduate Institute of Science and Technology

*Library: "a place set apart to contain books, periodicals, and other material
for reading, viewing, listening, study, or reference… (Merriam-Webster Inc, 1996),*

No longer "a place set apart," digital libraries offer an unprecedented opportunity to provide access to information that is integrated into work processes rather than separated from them. At the same time, there is an unprecedented potential for overwhelming users with excessive or irrelevant information, impairing their performance rather than improving it. Thus with the opportunity to create new models of what a library is and how it can be used comes the challenge of developing better understanding of the intended users, their work, and the circumstances under which they perform it.

"Tracking Footprints Through an Information Space," funded through the National Science Foundation Digital Libraries Initiative Phase 2, is a multidisciplinary collaboration among computer scientists and medical information professionals (Gorman, Delcambre, & Maier, 1998). The aim of this project is to develop tools that will help experts manage information in an increasingly digital workplace. We are particularly interested in learning how experts select and use information to solve complex problems so that we can develop tools to help them make better use of digital resources in their work.

In this paper we offer a brief overview of our experience using a variety of observational methods to better understand one class of users, expert clinicians treating patients in hospital settings. We show how our understanding of the users and their information management tasks has evolved based on observations in the field made by a multidisciplinary research team, and how this evolving understanding is guiding our efforts to create digital library technology. More complete descriptions of the background, methods, and findings of this work have been published elsewhere (Bowers & Delcambre, 2000; Delcambre, Gorman, Maier, Reddy, & Rehfuss, 1998; Delcambre et al., 2001; Gorman et al., 2000).

INITIAL APPROACH: USING PRÉCIS TO ASSIST FAMILIARIZATION

The Model

Our initial efforts focused on the problem of _familiarization_. We defined familiarization as the process whereby a clinician learns enough about the patient to feel comfortable making management decisions. The user in this model is a physician who must respond to a clinical

problem in a new patient, one whose complex medical history, multiple medical problems, and numerous medications are unfamiliar to the physician. The information task in this case is to become sufficiently familiar with the patient, their current condition, coexisting conditions, and current medications to determine a proper management plan. To accomplish this familiarization task a clinician must navigate through a complex collection of documents, identifying information that is relevant to the problem at hand, while ignoring the rest.

Anecdotal information and published reports (Nygren & Henriksson, 1992) suggested that clinicians use a variety of physical cues (size, thickness, color, wear-and-tear of pages) as they browse through paper-based medical records. These physical cues, however, are largely absent in digital collections such as the electronic medical record. To assist this browsing and selecting process and thereby facilitate the familiarization process in the expanded and distributed environment of an electronic medical record, we proposed to develop an application of superimposed information we called *précis*, a concise set of document descriptors that would, like the physical cues in a printed document collection, help a clinician estimate the relative utility of any document in the medical record without having to examine its content. Précis could thereby assist familiarization by facilitating more efficient navigation in a digital information space (Delcambre et al., 1998).

Observations in Think Aloud Sessions

To learn more about the user and the task, we conducted a small series of "think aloud" sessions (Ericsson & Simon, 1993) in which physicians were asked to use the medical record to solve clinical problems presented in a series of typical clinical scenarios prepared by the physician member of our research team (PG). We expected that the physicians would gather considerable amounts of background information to gain overall familiarity with each case, in addition to seeking out information specific to the presented problem. Instead, we found after just six of these sessions that physicians were much more focused in their information seeking, attending only to information that had direct bearing on the problem presented in the scenario. They spent little time on familiarization with the patient record as a whole, except in the case of a scenario in which they were to assume responsibility for ongoing, primary care of the patient.

Regardless of whether physicians had a near-term or long-term focus, we found that they invested considerable time, attention, and expertise in selecting which documents in the record to examine and which documents to ignore. In the process, we noted that physicians would often save their places in the record, marking the location of documents with their fingers as they continued to search the collection. In Figure 1, a physician has marked multiple locations of interest as she browses a medical record. Importantly, the decision to select a document for more careful examination was often based not on its content, but rather on its appearance, location in the collection, or other physical cues, a finding also reported by Nygren, et al. (Nygren & Henriksson, 1992).

Figure 1. Physician browsing through paper and electronic collections.

A second important observation in these sessions was that physicians frequently performed informal note taking and annotation in the process of selecting and retrieving information. These 'back-of-the-envelope' creations might be saved for later use or discarded. In either case, this behavior was observed repeatedly in our think aloud sessions, suggesting that this activity may play an important role in assembling and organizing relevant information to solve a clinical problem.

A third interesting observation during these sessions was actually a remark made by one of the physicians, to the effect that (paraphrasing), "It would be interesting to know what parts of the record Dr. S. (a cardiologist) looked at when she did her consultation on this patient." This suggested that the selections of one clinician might be useful to another, and led to the next phase of the work, focused on the *trace* left by an expert as she explores the collection of documents.

REVISED APPROACH: CAPTURING TRACES

The Model

Although limited in number, these initial observations suggested that familiarization, although it does occur, may be less important than we initially expected. Based on these observations we

refined our model of the user and the task, and revised our development approach, focusing on the problem of capturing the *trace* left by an expert as she examines the medical record. We observed that exploring a large complex document collection requires explicit choices about which documents to examine carefully and which documents to ignore. These choices require expert domain knowledge as well as knowledge about the format and organization of the documents and the collection. For a large collection, substantial time and attention had to be invested by an expert to locate a relatively small subset of relevant information. We speculated that these information selections would be of interest to subsequent users of the collection who were concerned with the same or a similar problem.

Figure 2. Illustration of "Traces" Experts Might Leave in Medical Record.

Taken together, these choices define a discrete subset of documents relevant to a given problem. This subset can become the beginning of what amounts to a thread or *trace* through the collection, evolving over time as the patient's condition evolves, attended to by various clinicians in a particular domain (for example, cardiology). With complex patients who have problems in multiple domains, many specialists may care for the patient, each creating a unique subset of the record that pertains to the clinical problem in their specialty. Thus multiple clinicians caring for a single patient may follow independent traces through the collection, each trace being a domain-specific or problem-specific subset of the medical record. Figure 2 illustrates this idea of independent but intersecting traces that might be left by experts in different

domains as they explore a collection in search of information relevant to the respective problems they are managing.

In this view of expert information seeking, the trace left by an expert through a collection of documents is an explicit set of statements about relatedness of those documents to one another and to the problem being addressed. That is, the domain knowledge of the expert is inherent in the subset of documents that she selects. With this in mind, we revised our development approach and chose to focus on capturing the trace left by the information seeking activity of an expert and to exploit the knowledge inherent in this trace to assist others who are concerned with the same problem.

The goal was thus to take advantage of the expert's knowledge rather than try to replicate it. In this way we hoped to reuse the attention, time, and expertise that is invested when an expert explores a document collection. The relatedness and usefulness of documents or items of information are thus defined by a domain expert in a specific context, rather than by any *a priori* context independent classification. To accomplish this we proposed to develop: (1) for each document, a précis of information about its origin, its content, and the history of its use by experts; (2) for each expert problem solver, a trace that describes the path taken through the collection; and (3) navigation tools that could assist subsequent problem solvers using the collection by exploiting the knowledge inherent in existing traces (for a more complete description see Gorman et al., 1998).

Key assumptions for this work were 1) that we could effectively capture this trace of information through diverse, complex document collections; 2) that the trace thus created could be reused at a later time by the same expert; and 3) that the trace could be reused by others concerned with the same problem. To test these assumptions, we returned to observation of clinicians, this time in the field rather than in the laboratory.

Pilot Observations in Healthcare Settings

With respect to the first assumption, pilot observations in multiple healthcare settings revealed a significant barrier to capturing and exploiting these traces: the number, diversity, and physical separation of the information systems that clinicians routinely employ. Following one emergency physician during routine work, we observed the use of some twenty separate systems and devices for storing, retrieving, and communicating information, including computer systems, paper records, and telecommunications technology, a finding confirmed by other investigators in this field (Enrico Coiera, personal communication). A further complication was the fact that these systems had each been designed to interact with a human user, but often had not been designed to interact with one another. To be useful, any system for capturing and tracking information selections in these settings would require technology that could effectively interact with multiple, disparate, distributed technologies.

Ethnography in an Intensive Care Unit
To test the second and third assumptions, that clinicians would reuse the subset of information previously created by themselves or others, we conducted ethnographic observation of nurses, physicians, and other clinicians in an intensive care unit. The observation and analysis team included members with diverse backgrounds to minimize bias in recording and interpreting

observations. Observations were conducted from December 1999 through February 2000, taking care to establish credibility and earn the trust of the staff to help improve the validity of the data. Findings from traditional participant-observation were enriched with interviews with key informants as well as with photographs and audio recordings. Findings were challenged, verified, and augmented through discussion with informants in focus groups, stimulated by examining images recorded in the course of fieldwork.[1]

We used an open-ended research question for this ethnography: "How do clinicians use information to help patients?" This helped to maintain a broad focus and enabled us to catalog the diversity of tools being used for recording and processing information, the many channels used to communicate information to others (Ames, 1993), and the multiple processes in place for sharing and making sense out of the information, individually and in groups. An important benefit of this observational work was an enriched view of information management behaviors in their natural context.

Observation in the ICU confirmed our assumptions about reuse of information selections. We observed many examples of experts selecting, organizing, and sharing subsets of relevant information, usually drawn from diverse, physically separate information sources (see Figure 3). Sometimes, this occurred through the use of formal documents that were part of the medical record, but often it took the form of informal, 'back of the envelope' creations, like the note-

Figure 3. Assembling a *bundle* of information from multiple sources.

---

[1] Visit http://medir.ohsu.edu/~gormanp/dli2/vr.html for examples of the images.

taking and annotation we had observed in our laboratory "think aloud" sessions (Marshall, 1998). In either case, we found that these collections of highly selected, organized, and annotated information, which we called *bundles,* were being used and shared regularly by individuals and groups to help them solve clinical problems and maintain individual and collective situation awareness (Gorman et al., 2000).

VERSION THREE: FACILITATING BUNDLING

In our observations in the intensive care unit, we were struck by how widespread was the creation, use, and reuse of these information bundles. Once again, observations from fieldwork enabled us to refine our understanding of the users, their tasks, and the tools they used to perform these tasks. This in turn led to a shift in the direction of our development team, leading to our current efforts toward developing technology that can facilitate the creation, use, and reuse of bundles. As a first step, we have developed a prototype application we call SLIMPad[2] (Delcambre et al., 2001).

Bundles: The Model

A more complete discussion of the bundles, their properties, creation, and use, and some of the challenges they present to digital libraries, has been published elsewhere (Gorman et al., 2000). A summary of the important properties of bundles are included in the Table below.

Table : Summary of Properties of Bundles

Bundles are Always	Bundles are Sometimes	Bundles are Not
actively created	organized	distributed
physical	dynamic	comprehensive
collections	temporary	simple lists
selective	diverse	forms or views
multigranular	complex	
context specific	multiauthored	
task oriented	shared	
redundant	multipurpose	
	uncertain	
	visual	

See (Gorman et al., 2000) for detailed discussion.

---

[2] Visit http://www.cse.ogi.edu/footprints/slimpad.php to view a mockup of SLIMPad.

Our fieldwork provided a number of additional observations which are important to an accurate understanding of these users and their tasks, and which must be kept in mind in developing any information tool meant to help them. These include:

- **Collaboration**: The "user model" for this work is most often not a single person but a group. Members of the group had roles, including information management roles that, although professionally and formally defined, are also dynamic and socially negotiated (Gorman, Lavelle, & Ash, 2002 (submitted)). Bundles were used in the ICU in a manner that fostered multi-author, multi-user collaboration.

- **Multithreading**: There is not one task but many, and they must often be completed in multiple stages. Frequent interruptions and simultaneous competing demands are constants in the critical care environment and to a lesser extent in other clinical activities. Bundles, as they are used by clinicians, appear to be an effective means of establishing and re-establishing situation awareness for an individual or a group, and facilitate resumption or transfer of interrupted tasks.

- **Physicality**: Critical care clinicians are certainly knowledge workers but they are first and foremost caregivers whose work is accomplished at the bedside. Although expertise dependent and information intensive, this work is undeniably physical in nature. The physical requirements of the work often preclude or interfere with use of existing technologies for data entry (mouse, keyboard, light pen, etc) and display (CRT or even LCD monitors). We observed the use of bundles with physical properties that permitted tight integration of the information tasks with the physical and mobile nature of the clinical tasks they support (Furuta, Marshall, Shipman, & Leggett, 1996).

- **Informality and flexibility**: Information in clinical work can be dynamic, uncertain, sensitive, or highly context dependent. The diagnosis and the management plan for a patient may actually be an evolving, socially constructed understanding. Formal documentation systems that require precise, explicit expression of relationships, categories, and interpretations may unnecessarily increase an already demanding cognitive load (Shipman & Marshall, 1999). We observed the use of bundles, sometimes temporary or in pencil, which permitted highly flexible format and content, enabling rather than inhibiting experts from adding meaning through organization and annotation of information.

Bundles: Focus Group Feedback

To verify our observations about bundles and evaluate the potential usefulness of the prototype application SLIMPad, we conducted focus groups with cardiologists and critical care nurses. Cardiologists responded enthusiastically to the SLIMPad application, in particular to its ability to reduce and organize information while preserving links to related underlying data. They expressed interest in seeing how this application might be further developed for use in transfer of care situations such as "handover", when one physician takes over for others on weekends. Critical care nurses responded positively to observations about bundles currently in use, which were predominantly paper-based, but appeared less enthusiastic about the prototype application, SLIMPad. We attribute this difference in response to two issues. First, the prototype application

was modeled after a physician-created bundle, possibly explaining the greater usefulness perceived by physicians. Second, nurses reemphasized the importance of portability and flexibility, because their work is always so directly physical – at the bedside. During focus groups it became apparent that the constraints of currently available technology for recording and display of information limit their usefulness in many bedside tasks.

Our current plans include further examination of specific tasks, as suggested by the physician informants, to which this technology may be adaptable. We believe a promising approach is videotaping of clinicians during handover and other activities to permit more precise analysis and comparison of task performance (Mackenzie, Hu, & Horst, 1995).

CHALLENGES

The approaches we have described here are associated with some important limitations and challenges. First, the sample sizes are in some cases quite small, potentially limiting reliability and generalizability of the findings. To reach firm and generalizable conclusions, findings would need to be confirmed in larger, systematically selected samples.   However, as Chelimsky has emphasized, when research and evaluation are conducted for the purpose of development or improvement, as opposed to producing generalizable knowledge, the design, methods, and relationship of evaluators to developers are quite different (Chelimsky & Shadish, 1997). Hence, as with techniques such as discount usability engineering (Nielsen, 1994), smaller sample sizes may be preferable when the objective is iterative refinement of a model or an application.

A second limitation is that our research to date has been confined to the health care domain, so it is not clear that these findings apply to expert problem solvers in other domains. Feedback from investigators in other domains to presentation of our findings (Gorman, Delcambre, & Maier, 2000) suggests that these findings may apply to expert problem solvers in other areas, such as aerospace and aviation. Our current plans include extending our observational methods to these other areas to determine whether this is so.

A third limitation is the jack-of-all trades problem: the challenge of developing sufficient methodological expertise to apply a wide variety of approaches to properly understand the work and the tools of the intended beneficiaries of the information technology under development. On the one hand, each new method requires acquisition of significant expertise if it is to be applied effectively to produce reliable results, leading to methodological specialization. On the other hand, specialization can lead to a sort of methodological tunnel vision, potentially limiting the ability to recognize and understand phenomena for which a particular method is not well suited. Ideally, these opposing constraints can be addressed through multidisciplinary collaboration, which combines intimate knowledge of the application domain, scientific expertise in the relevant computer and information sciences, and a methodological expertise that provides a repertoire of complementary investigative approaches.

CONCLUSION

This paper has briefly described how observation of clinical experts in the field has enabled us to improve our understanding of the intended users and the tasks they perform. Key elements, we believe, have been team composition, multimethod approach, and a relatively tight observation-development cycle. The multidisciplinary composition of our team has enriched our observations and improved the validity of our analysis and interpretations. The multiple observation methods we have employed, including "think aloud" clinical scenarios in the laboratory, participant observation in the ICU, key informant interviews, and focus groups have enabled us to enrich the data set, gain greater insight, and verify findings with informants. The relatively tight cycle of observation, analysis, development, and repeat observation has enabled us to iteratively and more rapidly refine our "user model" and "task model," improving, we hope, the usefulness of the technologies we are developing.

ACKNOWLEDGEMENTS

REFERENCES

Ames, S. W. (1993). *Multiple spoken and written channels of communication: an ethnography of a medical unit in a general hospital (Dissertation).* State University of New York at Buffalo, Buffalo, New York.

Bowers, S., & Delcambre, L. (September 2000). *Representing and Transforming Model-based Information.* Paper presented at the Workshop on Semantic Web, Held in Conjunction with the European Conference on Digital Libraries, Lisbon, Portugal.

Chelimsky, E., & Shadish, W. R. (Eds.). (1997). *Evaluation for the 21st Century.* Thousand Oaks, CA: Sage.

Delcambre, L., Gorman, P. N., Maier, D., Reddy, R., & Rehfuss, S. (1998). *Precis-based navigation for familiarization.* Paper presented at the MEDINFO 98 Proceedings of the Eighth World Congress on Medical Informatics [CDROM Proceedings], Seoul, Korea.

Delcambre, L., Maier, D., Bowers, S., Deng, L., Weaver, M., Gorman, P., Ash, J., Lavelle, M., & Lyman, J. (2001, April 2-6). *Bundles in Captivity: An Application of Superimposed Information.* Paper presented at the 17th International Conference on Data Engineering, Heidelberg.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol Analysis: Verbal Reports as Data.* Cambridge, MA: MIT Press.

Furuta, R., Marshall, C. C., Shipman, F. M., & Leggett, J. J. (1996, March 20-23). *Physical Objects in the Digital Library.* Paper presented at the Proceedings of the 1st ACM International Conference on Digital Libraries, Bethesda, MD.

Gorman, P., Ash, J., Lavelle, M., Lyman, J., Delcambre, L., & Maier, D. (2000). Bundles in the wild: Managing information to solve problems and maintain situation awareness. *Library Trends, 49*(2).

Gorman, P., Lavelle, M., & Ash, J. S. (2002, submitted). Order Creation and Communication in Healthcare. *Methods of Information in Medicine*.

Gorman, P. N., Delcambre, L., & Maier, D. (1998). *Tracking Footprints through an Information Space: Leveraging the Document Selections of Expert Problem Solvers (proposal).* National Science Foundation DLI. Available: http://www.dli2.nsf.gov/projects/ohsu.pdf [2001, 25 September].

Gorman, P. N., Delcambre, L. D., & Maier, D. (2000). *Tracking Footprint through a Medical Information Space.* Paper presented at the SmartSystems 2000: The International Conference for Smart Systems and Robotics for Medicine and Space Applications, Houston.

Mackenzie, C. F., Hu, P. F., & Horst, R. L. (1995). An audio-video system for automated data acquisition in the clinical environment. LOTAS Group. *Journal of Clinical Monitoring, 11*(5), 335-341.

Marshall, C. C. (1998, March 24). *The future of annotation in a digital (paper) world.* Paper presented at the 35th Annual GSLIS Clinic: Successes and Failures of Digital Libraries, University of Illinois at Urbana Champaign.

Merriam-Webster Inc. (1996). *Webster's Encyclopedic Unabridged Dictionary of the English Language*. New York: Gramercy Books.

Nielsen, J. (1994). Guerrilla HCI: Using Discount Usability Engineering to Penetrate the Intimidation Barrier. In R. G. Bias & D. J. Mayhew (Eds.), *Cost-Justifying Usability*. New York: Academic Press.

Nygren, E., & Henriksson, P. (1992). Reading the medical record. I. Analysis of physicians' ways of reading the medical record. *Computer Methods & Programs in Biomedicine, 39*(1-2), 1-12.

Shipman, F. M., & Marshall, C. C. (1999). Formality considered harmful: Experiences, emerging themes, and directions on the use of formal representations in interactive systems. *Computer Supported Cooperative Work, 8*(4), 333-352.

Scenarios in the Afya Project as a
Participatory Action Research (PAR) Tool for
Studying Information Seeking and Use
Across the Digital Divide

Bharat Mehra*, Ann Peterson Bishop*, Imani Bazzell**, Cynthia Smith*
* University of Illinois at Urbana-Champaign, ** SisterNet

INTRODUCTION

In the context of the "digital divide" that exists between information haves and have-nots, the Afya project takes a participatory action research (PAR) approach to improving access to health information and services for Black women. This initiative has involved forging community-level partnerships with SisterNet, a local grassroots group of Black women devoted to improving their physical, emotional, intellectual and spiritual health. Scenarios (or use-oriented design representations) have been significant in documenting information seeking and use in the context of health-related social practices as well as the broader socio-cultural context of users belonging to this traditionally marginalized group. Scenarios have been effective also as a PAR tool for establishing community-wide collaborations with Black women in the development of social and digital technologies associated with health information services.

Pointing to needed action for social justice, scenarios in PAR provide vignettes of information seeking and use grounded in harsh social realities experienced by Black women. Thus, they accentuate the urgency of achieving constructive social change based on redefined relationships. Needed changes involve strengthening community ties based on democratic practices that empower and engage Black women. In the context of community healthcare, scenarios as a socially grounded planning and design methodology have taken the form of personal narratives of Black women that capture social experiences and typical problematic health situations. The Afya project's technical agenda is geared towards building health services and information resources that are more usable and useful to Black women. Our social agenda involves improving interactions and outcomes in community health.

PAR involves people who are supposed to be beneficiaries of the research in all the stages of the research (Hall, et al., 1982). It has its roots in Paulo Freire's (1972) theory of conscientization, where he advocates a strategy for liberation of oppressed peoples via engagement and collaborations with local people for cultural action to oppose dominating power or cultural forces. PAR is thus an effective strategy for gaining citizen participation in the organization of local resources and community development (Chavis and Wandersman, 1990) that we believe can be valuable in the design and evaluation of information systems. PAR strives to help community members "validate and reframe information provided by their own life experience to enable them to take control of their surroundings and better determine their future" (Plaut, et al., 1992, p. 57). There is a close link between community development and PAR because both believe that the democratic values and processes that foster active and equitable citizens involvement are necessary for its success (Anyanwu, 1988). The Afya project taps the potential

of this connection between community development and PAR in the context of inequities in both health status and the use of information systems. Within the basic framework of PAR, scenario-building is helping us to carry out an action-oriented agenda for the empowerment of Black women.

SCENARIOS IN SOCIALLY GROUNDED RESEARCH

Within the past two decades, scenarios have acquired significant attention in theoretical and applied systems research, engineering practice, and in the discipline of human-computer interaction (HCI) studies. In its use of scenarios, the Afya project builds on previous research. At the same time, scenario building as a research method in the Afya project takes on new forms of expression, application, and use that are shaped by the tenets of PAR. The role of scenarios in the Afya project to 'map' values, practices, and the social context of users traditionally belonging to an underserved minority and marginalized group is markedly different from how scenarios have been employed in information systems research in the past. Afya's employment of scenarios is also different from traditional user-centered approaches to information-seeking and use research in its emphasis on the active participation and capacity-building of study "subjects" and the immediate improvement of local conditions for them. In the Afya project, local Black women no longer remain subjects in the traditional sense of the word. They are actively involved in describing the existing contextual realities of the situation, setting goals for improving the state of those conditions, and providing strategies for implementing changes. These are effectively practiced via the conceptual framework of scenario building that empowers community members themselves to describe, evaluate, analyze, interpret, and act via self-defined strategies.

In HCI, scenario-based design methodology—the "use-oriented perspective on the design and development of computer systems" (Carroll, 1995a, p. 2)—presents itself as a viable methodological tool for studying human-computer interactions. Scenario-based design methodology outlines a task-artifact framework based on an understanding that "system development starts from an analysis of implicit requirements embodied in user tasks" (Chin et al., 1997, p. 162). Scenarios are "artifacts in human activity" (Carroll, 1995b, p. 128) that take the form of concrete descriptions that record representative user goals and activities early and continuously in the system lifecycle. They attempt to capture details manifested in situations of use that overwhelmingly determine usability (Suchman, 1987).

Scenarios, thus, improve our skills in modeling user behavior in system design and development (Young and Barnard, 1991). The Alexandria Digital Library (ADL) project employed an analysis of information need and use scenarios related to georeferenced spatial data to design and evaluate digital information resources and services. In the Alexandria Digital Earth Prototype (ADEPT) project, a comprehensive follow-up on the ADL, scenarios are being used to build and assess spatial information service layers for users in university learning environments. (For more details, see Borgman et al., 2000, and the ADEPT website at http://alexandria.sdc.ucsb.edu/adept/adept.html). Such scenario-based studies in HCI have contributed to information systems research by outlining theoretical and practical design frameworks that explore the interactions between humans and computer systems. They present mechanisms for modeling user behavior and providing useful feedback about usability requirements of end users. The central focus for the use of scenarios in these spheres of

application has been towards the design and development of systems that can cope with users' changing demands and needs. There have been other useful methodologies employed in HCI that are similar to scenario building. Participatory design involves users directly in the design process and aims at improvements in their work conditions (see Kyng, 1995; Muller, Kuhn, Wildman, & White, 1993; Schuler and Namioka, 1992). Contextual design uses ethnographic techniques for understanding human behavior related to computer use (Whiteside, Bennett, and Holtzblatt, 1988; Wixon, Holtzblatt, and Knox, 1990; Holtzblatt and Beyer, 1993).

PAR has a strong basis in motivating citizen involvement for empowerment through "the process of constructing and using their own knowledge" (Reason, 1994, p. 328). Scenarios allow an understanding of how people actually use technology and the descriptions of their experiences to shape outcomes in the design of computerized information systems. Ethnographic methods immerse the researcher in actual settings to study how participants behave and act in real-life situations. In traditional approaches from different disciplines, these three methodological strategies—PAR, scenario building, and ethnographic study—have often been unrelated and isolated in their applications. For example, participatory design may not necessarily involve the use of scenarios; scenarios may be written and applied towards system design with no participation from users; and even ethnographies do not always involve study participants in interpretation and analysis of the results[1]. The Afya project attempts to interweave these three research methodologies in the context of the use of technology to promote social justice and empowerment of local Black women. It promotes the intersection of PAR, scenarios, and ethnographic observation by redefining relationships to balance existing power dynamics in the local health information service environment.

Hence, a major point of variation in the use of scenarios in the Afya project has been its explicit, much broader, aim of increasing social justice at the level of the community as a whole. Scenarios in the Afya project do not so much express information behaviors and social practices of users in isolated contexts of a particular information system as give voice to the socio-cultural context of a user community traditionally excluded in the system design process. They express the multiple nuances of marginalized experiences as embedded within the larger socio-cultural context of the disfranchised community. Though scenarios are applied towards achieving social and technical outcomes of change in the health information and service environment, the goal is not just that. Afya's vision is to empower Black women to take constructive action for social equity and social justice at the community level, an agenda that encompasses many dimensions that go beyond mere health and information issues. In order to provide a collaborative milieu of interaction between Black women, academic professionals, and health care and information providers, Afya's scenario-building aims to provide a common vocabulary for giving expression to this vision, and become an action-oriented tool for facilitation of this goal. Analysis of Black women's personal experiences narrated in their scenarios sheds light on the nature of the socio-political and socio-economic realities of our local community, on how, these are embodied in the use and impact of health information services, and on how Black women can begin to create a new reality for themselves.

---

[1] As pointed out by an anonymous reviewer of this paper.

71

SCENARIOS IN THE AFYA PROJECT

The Afya project attempts to foster social justice by nurturing equitable and participative social activities around technological development and use. In contrast to the goal of scenario building in most HCI studies, the design of an information system is not the main goal of the Afya project. We view technology as only one of the tools for addressing deeper socio-cultural realities. In this context, the use of scenarios in the Afya project has been to establish a *communication and information exchange* about the complex nature of some of the harsh realities surrounding the provision of community health and information services for Black women. We are also employing scenarios to formulate and carry out an *action-oriented agenda* that supports Black women in realizing their own potential in promoting social justice.

In supporting such community-based participatory action research endeavors, scenarios are Black women's expressions of specific situations they experience in health-related environments. Scenarios incorporate local Black women's goals, perceptions, activities, barriers, and other socio-cultural, economic, political, and environmental factors that influence their use of existing health-related information services and shape their information seeking practices. Use scenarios holistically encapsulate Black women's values, social relationships, and experiences. They also point to action-oriented strategies that women themselves outline and can implement for initiating the equitable social and technological distribution of health care and information resources. Two major scenario-building initiatives conducted in the Afya project were their development during focus group discussions and their application in situated user evaluation workshops (For a detailed description of scenario methodology in the Afya project, see Bishop, et al., in press, a):

Data gathered during four focus groups conducted in January and February 2000 provided descriptions of health-related information seeking behaviors of SisterNet women. The scenarios narrated by local Black women also revealed the social context and associated relationships in community healthcare information services. Imani Bazzell, SisterNet Director, served as focus group moderator. Scenarios were generated by discussing: participants' vision of a healthy Black women; significant health concerns or situations they recently experienced; resources for getting health information; positive and negative factors in obtaining health information and services; use of computers; and actions participants could take to improve health and information services.

The topics discussed in the focus groups holistically address authentic marginalized experiences of local Black women in the health information service environment because they intrinsically link information technology to context, relationships, and practices. The discussions formed the backbone for use scenarios that are now being applied in the design of community-based online health information services. Local Black women have taken the lead in analyzing their own scenarios in the course of evaluating Internet-based health information and communication services in order to build appropriate web-based tools that are directly relevant to their needs and expectations.

The three focus groups were conducted with 6-8 SisterNet women, who, based on their age, were assigned into either "Young Women's," "Prime Time," or "Wise Women" group. In addition, one focus group was conducted with participants who were community health care and health

information providers. Scenarios were effective in highlighting key elements related to community health care and information provision, including:

- The embeddedness of health information services within a larger socio-cultural and environmental context;

- The role of political, economic, psychological, gender and racial issues, as well as the lack of available technological and information infrastructure, in shaping access and provision to health care and information delivery for local Black women;

- The importance of strengthening collaborative, community wide efforts;

- A need for redefining of relationships that will contribute towards scaling of traditional power dynamics between Black women and health information service providers; and

- Possible participative, action-oriented strategies that could lead towards social empowerment of Black women.

Application of scenarios from the focus groups played an important role as a methodological tool in situated evaluation studies of online health information services. The following objectives guided the situated user evaluation workshops that were conducted in the Afya project:

- Engage SisterNet women in the process of evaluation of online health-related resources for Black women;

- Based on input from local Black women, provide constructive critiques of online health-related resources created by local librarians and health information and service providers;

- Increase local Black women's interest and ability in use of computer technology, the Internet and Prairienet;

- Develop a template for a prototype web site for SisterNet and elicit direct feedback from SisterNet members in evaluation of the site.

Black women's scenarios specifically attempted to bridge online and offline worlds by grounding the assessment and evaluation of online services in the lifeworlds of local SisterNet women. Online services such as local hospital websites and the health and human service directory of Prairienet (the local community network in East Central Illinois), national health portals, and electronic discussion groups related to Black women's health were evaluated by SisterNet women according to how well they addressed the situations captured in their scenarios. Situated user evaluation workshops also helped elicit new scenarios from the local Black women participating at the evaluation workshops, who expressed their own personal experiences.

During spring 2000, a comprehensive set of scenarios was developed from transcripts and notes taken during the focus groups and the situated evaluation workshops. Initially, it was not clear what a scenario would look like, and unsuccessful attempts were made to classify scenarios with categories like topic, purpose, type, mode, query, qualities, format, user, etc. Once we began to analyze what local SisterNet women had to say about their marginalized experiences in the

health information service environment, we realized the inadequacy of this rigid fielded structure for expressing the varied qualitative data of experiential narratives. It was replaced by a more flexible understanding of scenarios, one that was developed in the form of documenting descriptive anecdotes of Black women's experiences in the health care information context. These were more apt in understanding what kind of strategies would be effective in bringing social change via social and digital technologies that would lead towards empowerment of SisterNet women.

The main challenge was to piece together the elements—key social realities, specific instances of local Black women's health experiences, and questions that mirrored information needs—in the form of scenarios that formed a coherent narrative embodying real life situations. The two examples presented below illustrate how scenarios were actually pieced together early in the Afya project. For the first scenario, three comments made by SisterNet women in focus groups were brought together; two were comments about stress, and one was related to a specific inquiry about mental health resources:

> "I am a grandmother and I think we find ourselves under so much more stress every day. I know I do. I am in a situation where, this past May I had to take two of my grandchildren rather than have them go into foster care so that is an added stress in my life. […] In fact, since I have had the two kids [one and two year olds], they bring home everything [i.e., illnesses] from daycare, you know. […] I have just really been sicker this year than ever. Mainly because of these two babies and the added stress […] (FG#2Ap.4). I am always bothered by little problems that seem to become much bigger in my mind. [I am] always complaining and feeling weak and insufficient to cope with my situation (FG#2Ap.1). {How can I increase my self-esteem, improve my state of mind, stop feeling depressed and start feeling better about myself?"} (FG#2Ap.2).

All three comments were found in the "Prime Time" Focus Group (FG#2). The location of the comments is recorded by the noting of the approximate page numbers (p.4; p.1; p.2) in the focus group notes from where they were taken. The use of the letter "A" identifies the original focus group notes as the source of the comments. Curly brackets are used to designate the motivating needs reflected in this scenario in the form of a question about information on relevant resources for mental health.

Similarly, the following scenario was created from the amalgamation of four comments made by SisterNet women: one comment about uterine cancer, one on its potential symptoms, one about a lack of money, and the last about a need for insurance coverage.

> "My mother had died of uterine cancer last year, and we did not even know till she died that she had such "female trouble" (FG#2Ap.3). Then just a month back, I myself suddenly developed terrible pains in my lower abdomen, and now I can feel a lump the size of a small walnut in that area (FG#2Ap.3). I don't have any money (FG#2Ap.7) to go regularly for checkups to my doctor and my insurance does not cover more than one trip per month. {Are there any insurance companies that will provide better deals in health care coverage and lower rates than the one I already have} (FG#1Ap.2).

The two comments about uterine cancer and its perceived symptoms are from the "Prime Time" Focus Group (FG#2), while the other two comments about lack of money and the insurance company query are from the "Young Women's" Focus Group (FG#1).

As may be obviously noted, the two scenarios described above are unauthentic owing to their derivation from focus group recorder notes; construction from the narration of several different women; and the articulation of questions that may have originally been explicit or implicitly stated in the SisterNet speaker's remarks. In the Afya project's situated user evaluation workshops, we intentionally explored such scenario constructions because linking various combinations of narrated experiences gave us an opportunity to represent a larger number of possible lived situations experienced by Black women. Different forms of scenarios, including verbatim transcripts of individual narratives, were used in different stages of our research.

In other words, the purpose of scenario building changed considerably during different stages of the Afya project. Some functions of our scenarios are listed below:

- Identified SisterNet women's needs and experiences;

- Targeted barriers faced by local Black women;

- Prompted other Black women to share their own health-related experiences;

- Applied in evaluation of online health information resources;

- Helped in assessment of real health information services such as those found at local health clinics, public libraries, etc.

- Provided a framing device for health care and information providers to collect and present appropriate health information resources;

- Became an instrument for organizing information in a SisterNet resource guide in a manner that would be relevant to community members;

- Contributed towards collection and presentation of suitable information for informational health profiles prepared by Black women; and

- Helped in the design of health-related web resources.

Both focus groups and user evaluation workshops provided an avenue for observing and *creating* social interactions across various community stakeholders because participants were brought together in relatively informal social situations that involved interactions within and among SisterNet women, information and health care professionals, and researchers. Scenarios elicited from local Black women during focus group discussions and situated user evaluation workshops captured vignettes of social realities that obviously needed improvement. Scenarios are helping us identify and negotiate *as a community* the larger political, economic, socio-cultural, and environmental factors affecting the provision of community health care and information services. Scenarios capture problems considered significant by local Black women and providers and their impact on health service and information delivery. These included problems in social interactions among local Black women and between women and health service providers as well as problems accessing and exchanging (both online and offline) information.

An important note that should be made here is concerning the situated nature of our methodology for expressing how scenarios were created and applied for different purposes. However, this does not imply that no other possible approaches should be tried in different circumstances; the richness of scenario application lies in developing different methodologies for creation of scenarios in different contexts, based on the situational dynamics, contextual goals and potential compliances, user needs, practices and behavior, researcher temperaments and preferences, etc.

AFYA'S ACTION-ORIENTED AGENDA
FOR THE EMPOWERMENT OF BLACK WOMEN

Scenarios narrated during focus groups and situated user evaluation workshops suggest bringing about social change by building upon the existing capacity of Black women. In their scenarios, Black women provided many examples of the need to take charge of their destinies via developing their own social, technical, and information resources. In the next stage of the Afya project, SisterNet women developed a Community Action Plan that encapsulates their vision of achieving social justice. The Community Action Plan brings together a wide range of strategies engaging local Black women in building democratic partnerships and nurturing their interest, proficiency, access, and participation related to health information and the Internet (Bishop et al., 2001).

The following section presents examples of how scenarios played a role in developing a Community Action Plan that acknowledges and addresses Black women's harsh experiences in health-related situations. These scenarios represent complex, interconnected and multidimensional realities surrounding the provision of health care and information services to Black women. The action-oriented strategies proposed by SisterNet women in their Community Action Plan suggest unique options for addressing the multidimensional realities in a manner that will contribute towards achieving social justice via the empowerment of Black women in the community.

Issues such as institutionalized prejudices, politicized behaviors based on racist and sexist policies, inequality and discrimination often came up in local Black women's narration of their health-related social experiences. For example, one young woman said:

> *"It's hard to know if it is really racism or if the health care providers are being pushed. The fact that if a Black woman presents with abdominal pains, the first thing they want to do is run a series of venereal disease tests on her; whether that is the issue or not."*

This scenario expresses mistrust, based on gender and racial prejudices, which has led to poor provider-patient relationships. One of the major activities proposed in Afya's Community Action Plan was the organization of a Black women's health fair designed to increase knowledge about health concerns and resources as well as provide an opportunity for positive social interaction between health information and service providers and community women.

During the SisterNet Health Fair held in the community room of one of our local public libraries in May 2001, the university health sciences librarian provided information aimed at encouraging Black women to use the campus collections. She developed welcoming brochures that

emphasized that library resources were freely available to all community members and identified health-related resources relevant to Black women that could be found at the library. Another important Health Fair goal was to build upon skills related to the use of the Internet for finding health information; the health sciences librarian also helped prepare handouts containing information about online resources that her staff were especially well equipped to offer to the general public. For example, instruction sheets for searching PubMed, the National Library of Medicine's (NLM) journal literature search system were made available with information about customer support and help sources, database coverage, tips on searching information, and examples of articles on Black women's health. Such interactions between local health-related information service providers and Black women in the community is contributing towards openness in communication channels between various stakeholders and leading toward the creation of more positive social environments for seeking health information.

Afya's scenarios also suggested a lack of culturally sensitive and relevant information as a significant concern, as in the following comment from a health provider:

> "We sometimes still have trouble educating professionals because materials still come from white male medical models."

Another significant cause for concern were descriptions of SisterNet women's perception of their lack of confidence in demanding good health services and information, e.g.:

> "We as women have been taught that this is the doctor and they know what's good for us and so we listen to the doctor. You don't challenge his word."

Afya's Community Action Plan proposes to improve the quantity and quality of health information from local providers. Additionally, and more significantly, the agenda aims to establish and institutionalize ongoing information provision from Black women, in both digital and print formats, as an important aspect of building their self confidence. Local Black women are developing health profiles (brief articles addressing common health concerns of Black women), a resource guide on health-related information and technology use, and a SisterNet website. The SisterNet website promises to feature jargon-free, culturally appropriate health information for promoting the physical, emotional, spiritual, and intellectual well being of SisterNet women. It will also contain culturally appropriate and relevant online communication spaces such as chatrooms, access to health-related local news, tips, and public policy information.

Afya's Community Action Plan is also involving SisterNet women in the development of appealing, easy to read health promotional materials that will be accessible in both print and online formats at sites such as popular hair salons, local churches, and women's health clinics. We propose developing five SisterNet Resource Centers in locations such as these that are comfortable and convenient for Black women. This will help in the diffusion of such materials and broaden outreach to community members who lack strong social networks. Providing Black women an opportunity to be involved in the creation of culturally appropriate resources is an important step for building self-esteem and equalizing the power imbalance between Black women and other community health stakeholders. Creating culturally appropriate resources also involves the representation of home remedies and "folk wisdom," a need expressed in scenarios

such as the following, which advocated the capture of the traditional local knowledge held by older Black women:

> *"Most women that are retired and at home, they have some kind of educational background. They're been out in the world. They have more "mother wit" than the average person with a Ph.D. I am concerned we are going to lose stuff as people die if we don't document it. We have got to write it down or it's going to be lost forever."*

Scenarios of SisterNet women also pointed towards social isolation as a major community problem that is associated with health. For example, several older African American women from the community complained of depression due in part to the lack of social interaction with other community members:

> *"The problem is that I have nobody to talk to and I really feel really alone. I would like somebody to do things for me to show that someone cares. I can't afford to hire people to do things for me like shovel snow, cut grass, and stuff like that, and this makes me very depressed."*

Afya's Community Action Plan is facilitating the development of a strong social network among SisterNet women for the exchange of support and information related to both health and computing. We hope that social interactions established offline and online will promote stronger community ties for sharing information and providing social support. Several older Black women who complained of loneliness noted that participation in chatrooms and listservs through the SisterNet website might be one way to provide themselves with more opportunities for social support and interactions with other Black women in the community. Younger women suggested that posting a SisterNet events calendar and providing an online means to find sports partners would increase their ability to create and participate in offline social activities related to pursuing a healthier lifestyle. The lack of computer skills described in a number of scenarios as a major barrier to accessing health information and services is also being addressed within the context of developing stronger social ties to support technology training. Under the auspices of Afya's Community Action Plan, small cohorts of SisterNet women have been provided with computers and training, in informal social settings, to overcome this lack of technology access and associated know-how. Constructive change is being further enacted as SisterNet women who have received computers and training serve as informal technology mentors to others in their personal social circles.

CONCLUSION

In the Afya project, we are exploring an approach to the study of information seeking and use that is grounded in the basic tenets of participatory action research. While LIS has a strong tradition of excellent research on the information needs and behavior of marginalized groups in society—embodied in the work of Brenda Dervin, Elfreda Chatman, and others—what's missing is a methodological stance that foregrounds the active participation of marginalized groups in all phases of research along with the achievement of social justice, capacity-building, and immediate improvements in local conditions (Bishop et al., in press, b). Within the basic framework of participatory action research, we have been experimenting with scenario building as a specific technique for gaining the active participation of Black women and representing their

lifeworlds and local knowledge throughout the process of developing online services related to community health.

The use of scenarios in the Afya project raises an awareness of some of the socio-cultural, economic, political, and environmental issues impacting provision of healthcare information and services in the community. This may ultimately contribute toward strengthening community ties and building social networks of interaction and exchange between local women, health service providers, and other stakeholders involved in the distribution of health information and services. Additionally, scenarios of Black women point towards an action-oriented agenda for addressing health concerns in the community. The scenarios developed in Afya's focus groups and applied in situated usability testing workshops suggest specific strategies for empowering Black women to help build social and digital technologies that we hope will make the provision of healthcare in our community more just.

REFERENCES

Anyanwu, C. (1988). The technique of participatory research in community development. *Community Development Journal, 23*, 1, 11-25.

Bishop, A., Mehra, B., Bazzell, I., & Smith, C. (in press, a). Scenarios in the design and evaluation of networked information services: An example from community health. In C. McClure & J. Bertot (Eds.), *Evaluating networked information services: Techniques, Policy, & Issues*. Information Today.

Bishop, A., Mehra, B., Bazzell, I., & Smith, C. (in press, b). Participatory action research and digital libraries: Reframing evaluation. In A. Bishop, N. A. Van House, and B. P. Buttenfield. *Digital library use: Social practice in design and evaluation.* Cambridge: MIT Press.

Bishop, A., Bazzell, I., Mehra, B., & Smith, C. (2001). Afya: Social and digital technologies that reach across the digital divide. *First Monday*, 6, 4. [http://www.firstmonday.org]

Borgman, C. L., Gilliland-Swetland, A. J., Leazer, G. H., Mayer, R., Gwynn, R., & Mautone, P. (2000). Evaluating digital libraries for teaching and learning in Undergraduate education: A case study of the Alexandria Digital Earth Prototype (ADEPT). *Library Trends, 49*, 2.

Carroll, J. (1995a). Introduction: The scenario perspective on system development. In J. Carroll (Ed*.), Scenario-based design: Envisioning work and technology in system development* (pp. 1-17). NY: John Wiley & Sons.

Carroll, J. (1995b). Artifacts and scenarios: an engineering approach. In A. Monk & N. Gilbert (Eds.), *Perspectives on HCI: Diverse approaches* (pp. 121-144). San Diego, CA: Academic Press.

Chavis, D., & Wandersman, A. (1990). Sense of community in the urban environment: A catalyst for participation and community development. *American Journal of Community Psychology,18*, 1, 55-81.

Chin, Jr., G., Rosson, M., & Carroll, J. (1997). Participatory analysis: Shared development of requirements from scenarios. In S. Pemberton (Ed.), *CHI 97 – Human factors in computing systems: Conference proceedings* (pp. 162-169). NY: ACM.

Freire, P. (1972). *Cultural Action for Freedom*. Cambridge: MA: Penguin Education.

Hall, B., Gillette, A., & Tandon, R. (editors). (1982*). Creating knowledge: A monopoly? Participatory research in development.* New Delhi: Society for Participatory Research in Asia.

Holtzblatt, K. & Beyer, H. (1993). Making customer-centered design work for teams. *Communications of the ACM, 36*, 92-103.

Kyng, M. (1995). Creating Contexts for Design. In J. Carroll (Ed.*), Scenario-based design: envisioning work and technology in system development* (pp. 85-107). NY: John Wiley & Sons.

Muller, M., Kuhn, S., Wildman, D., & White, E. (1993). Guest editors with taxonomy of participatory design practices: A brief practitioner's guide. *Communications of the ACM, 36*, 24-28.

Nardi, B. and O'Day, V. (1999). *Information ecologies: Using technology with heart.* MIT Press.

Plaut, T., Landis, S., & Trevor, J. (1992). Enhancing participatory research with the community oriented primary care model: A case study in community mobilization. *The American Sociologist/Winter,* 56-70.

Reason, P. (1994). Three approaches to participative inquiry. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of Qualitative Research* (pp. 324-339). New Delhi: Sage Publications.

Schuler, D., & Namioka, A. (1992). *Participatory design: Perspectives on system design.* Hillsdale, NJ: Lawrence Erlbaum.

Suchman, L. (1987). *Plans and situated actions: The problems of human-machine communication.* Cambridge: Cambridge University Press.

Whiteside, J., Bennett, J., & Holtzblatt, K. (1988). Usability engineering: Our experience and evolution. In M. Helander (Ed.), *Handbook of human-computer interaction* (pp. 791-818). Amsterdam: Elsevier Science.

Wixon, D., Holtzblatt, K., & Knox, S. (1990). Contextual Design: An emergent view of system design. *In Proceedings of CHI'90: Conference on Human Factors in Computing Systems, Seattle, April 1-5* (pp. 329-336). NY: ACM Press.

Young, R., & Barnard, P. (1991). Signature tasks and paradigm tasks: New wrinkles on the scenario methodology. In D. Diaper & N. Hammond (Eds.), *People and Computers VI: Proceedings of the HCI '91 Conference 20-23 August 1991* (91-101). Cambridge, U.K.: Cambridge University Press.